

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 09:54:25 ; Search time 987 Seconds
(without alignments)
7503.317 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggaggggggtccagctgt.....agggaagcaggctgcattcc 1515

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3216467 seqs, 2444149694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1515	100.0	2190	14	US-10-135-689-1
2	1515	100.0	2190	17	US-10-690-617-1
3	1510.2	99.7	1937	13	US-10-333-314-39
4	1507	99.5	2018	13	US-10-302-172-433
5	1505.4	99.4	1611	13	US-10-092-900A-55
6	1490.6	98.4	1683	9	US-09-922-138-9
7	1490.6	98.4	1686	17	US-10-391-364-43
8	1490.6	98.4	2711	9	US-09-922-138-7
9	1490.6	98.4	2711	17	US-10-391-364-41
10	1457.2	96.2	3501	13	US-10-168-582-14
11	1455.8	96.1	1547	13	US-10-092-900A-53
12	1381.4	91.2	1725	13	US-10-092-900A-57
13	651.2	43.0	4427	15	US-10-316-124-7
14	651.2	43.0	4427	15	US-10-316-124-8

15	651.2	43.0	4942	15	US-10-205-823-51	Sequence 51, Appl
16	651	43.0	1804	15	US-10-205-823-53	Sequence 53, Appl
17	651	43.0	5611	16	US-10-295-027-1138	Sequence 1138, Ap
18	643	42.4	2056	14	US-10-098-841-22	Sequence 23, Appl
19	529.6	35.0	705	9	US-09-764-868-225	Sequence 225, App
20	529.6	35.0	705	11	US-09-764-875-182	Sequence 182, App
21	529	34.9	731	13	US-10-296-115-20	Sequence 20, Appl
22	518	34.2	1927	14	US-10-098-841-23	Sequence 23, Appl
23	361	23.8	29629	14	US-10-135-689-3	Sequence 3, Appl
24	361	23.8	29629	17	US-10-690-617-3	Sequence 3, Appl
25	354.4	23.4	432	17	US-10-283-975A-847	Sequence 847, App
26	348.8	23.0	439	13	US-10-296-115-247	Sequence 247, App
27	272	17.0	876	13	US-09-823-245A-163	Sequence 163, App
28	262.6	17.3	491	10	US-09-918-995-1271	Sequence 1271, Ap
29	246.8	16.3	1652	13	US-10-425-114-26840	Sequence 26840, A
30	244.6	16.1	485	10	US-09-918-995-2780	Sequence 2780, Ap
31	201.2	13.3	469	9	US-09-864-761-5606	Sequence 5606, Ap
32	201	13.3	218	9	US-09-864-761-22375	Sequence 22375, A
33	142.8	9.4	288	16	US-10-305-720-1008	Sequence 1008, Ap
34	139	9.2	463	9	US-09-764-869-1603	Sequence 1603, Ap
35	139	9.2	463	15	US-10-091-504-1603	Sequence 1603, Ap
36	139	9.2	463	16	US-10-227-577-1603	Sequence 1603, Ap
37	136.8	9.0	463	9	US-09-764-869-192	Sequence 192, App
38	136.8	9.0	463	15	US-10-091-504-192	Sequence 192, App
39	136.8	9.0	463	16	US-10-227-577-192	Sequence 192, App
40	113.8	7.5	529	9	US-09-954-456-1333	Sequence 1333, Ap
41	113.8	7.5	529	9	US-09-880-107-1390	Sequence 1390, Ap
42	113.8	7.5	1743	13	US-10-425-114-15808	Sequence 15808, A
43	110.8	7.3	1268	13	US-10-425-114-33104	Sequence 33104, A
44	106	7.0	1448	17	US-10-437-963-59939	Sequence 59939, A
45	106	7.0	1933	13	US-10-425-114-3580	Sequence 3580, Ap

ALIGNMENTS

RESULT 1
US-10-135-689-1
; Sequence 1, Application US/10135689
; Publication No. US2002123121A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; PRIOR FILING DATE: 2002-05-01
; PRIOR FILING DATE: 2000-11-13
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-135-689-1

Query Match	100.0%;	Score	1515;	DB	14;	Length	2190;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1515;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGTAGAACGGGTG	60				
Db	173	ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGTAGAACGGGTG	232				
Qy	61	GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTCGCCAGAGCTTACTAGAAC	120				
Db	233	GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTCGCCAGAGCTTACTAGAAC	292				
Qy	121	GGTGTGGACCCCCACCCAGCGGCGGAGCTGCCTCTGTGATCCCTGTGGCAGTACTTCAAGA	180				

Db 293 GGTGTGACCCCCACACAGGGCGAGAGCTGCTCTCTGTGATCCCTGCGCAGTACTTCAAGA 352
QY 181 CTGCTCCAGCCCGGCCCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCAGCA 240
Db 353 CTGCTCCAGCCCGGCCCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCAGCA 412
QY 241 GGAAGCTATCTGAGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC 300
Db 413 GGAAGCTATCTGAGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC 472
QY 301 CCCCGGCCCTGGCGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCGAG 360
Db 473 CCCCGGCCCTGGCGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCGAG 532
QY 361 GACTGGCTGAGCTGAACCAAGTCAAGCTGCAGAGTGAGATGGCAAGGGTGGCTACGGT 420
Db 533 GACTGGCTGAGCTGAACCAAGTCAAGCTGCAGAGTGAGATGGCAAGGGTGGCTACGGT 592
QY 421 GTGCTGAGGCTGGCCTACAAGCAAGTGAAGACAGACACTATGCAATGAAGTCTTTTCC 480
Db 593 GTGCTGAGGCTGGCCTACAAGCAAGTGAAGACAGACACTATGCAATGAAGTCTTTTCC 652
QY 481 AAAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 540
Db 653 AAAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 712
QY 541 GCTGCCAGGAGGAGACCAAGCAGTCTGCTGCCCTGGAGCGGGTGTACCAAGGAGATT 600
Db 713 GCTGCCAGGAGGAGACCAAGCAGTCTGCTGCCCTGGAGCGGGTGTACCAAGGAGATT 772
QY 601 GCCATCTGAAGAGCTGGACAGCTGAATGTGGTCAAACTGATCGAGTCTTGGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACAGCTGAATGTGGTCAAACTGATCGAGTCTTGGATGAC 832
QY 661 CCAAGCTGAGACAACTCTATTTGGTGTGTGACTCTGAGAAAGGGCCCGCTCATGGAA 720
Db 833 CCAAGCTGAGACAACTCTATTTGGTGTGTGACTCTGAGAAAGGGCCCGCTCATGGAA 892
QY 721 GTGCCCTGTGAAGAAGCTTCTCGAGGAGCAAGTGGCTCTACCTGCGGGACGTCAATC 780
Db 893 GTGCCCTGTGAAGAAGCTTCTCGAGGAGCAAGTGGCTCTACCTGCGGGACGTCAATC 952
QY 781 CTGGGCTCGAGTACTGACATGCCAGAGATCGTCCAGAGCATCAAGCCATCCAAAC 840
Db 953 CTGGGCTCGAGTACTGACATGCCAGAGATCGTCCAGAGCATCAAGCCATCCAAAC 1012
QY 841 CTGCTCTGGGGATGATGGGCAAGTGAAGTCCCGACTTTGGCGCTCAAGCAACCAAGTTT 900
Db 1013 CTGCTCTGGGGATGATGGGCAAGTGAAGTCCCGACTTTGGCGCTCAAGCAACCAAGTTT 1072
QY 901 GAGGGAAACGACCTCAGCTGTCCAGACGGCGGGAACCCAGCATTCATGGCCCCCGAG 960
Db 1073 GAGGGAAACGACCTCAGCTGTCCAGACGGCGGGAACCCAGCATTCATGGCCCCCGAG 1132
QY 961 GCATTTCTGATTTCCGCGCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCACTGGC 1020
Db 1133 GCATTTCTGATTTCCGCGCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCACTGGC 1192
QY 1021 GTACGTTCTACTGCTTTTGTCTATGGGAAGTGCCTATTCATCGACGATTCATCTGGCC 1080
Db 1193 GTACGTTCTACTGCTTTTGTCTATGGGAAGTGCCTATTCATCGACGATTCATCTGGCC 1252
QY 1081 CTCCACAGGAAGTCAAGATGAGCCGCTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCCACAGGAAGTCAAGATGAGCCGCTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1312
QY 1141 GAGCTCAAGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGATTGGGGTG 1200
Db 1313 GAGCTCAAGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGATTGGGGTG 1372
QY 1201 CAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGAGGAGGCCCTTCTTCGGAG 1260

Db 1373 CCAGACATCAAGTTGCAACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTTCGGAG 1432
QY 1261 GAGGAGCACTGCAGCGTGGTGGAGGTGACAGAGGGGAGGTAAAGAACTCAGTCAAGGCTC 1320
Db 1433 GAGGAGCACTGCAGCGTGGTGGAGGTGACAGAGGGGAGGTAAAGAACTCAGTCAAGGCTC 1492
QY 1321 ATCCAGCTGGAACCAAGTGCATCCCTGGTGAAGTCCATGTCTGAGGAAAGCTTCTTTGGG 1380
Db 1493 ATCCAGCTGGAACCAAGTGCATCCCTGGTGAAGTCCATGTCTGAGGAAAGCTTCTTTGGG 1552
QY 1381 AACCCGTTTGGAGCCCGAGGCAAGGAGGAGAGGATCCATGTCTGTCAGGAAACCTA 1440
Db 1553 AACCCGTTTGGAGCCCGAGGCAAGGAGGAGAGGATCCATGTCTGTCAGGAAACCTA 1612
QY 1441 CTGCTGAAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1500
Db 1613 CTGCTGAAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687

RESULT 2
US-10-690-617-1
; Sequence 1, Application US/10690617
; Publication No. US2004008926A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904DIV II
; CURRENT APPLICATION NUMBER: US/10/690,617
; PRIOR FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-690-617-1

Query Match 100.0%; Score 1515; DB 17; Length 2190;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGGGGTCCAGCTCTCTGCTCCAGGATCTCTCGGCGAGAGCTGTGTAGAACGGGTG 60
Db 173 ATGAGGGGGTCCAGCTCTCTGCTCCAGGATCTCTCGGCGAGAGCTGTGTAGAACGGGTG 232
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTGGCCAGAGCTTACTAGAAC 120
Db 233 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTGGCCAGAGCTTACTAGAAC 292
QY 121 GGTGTGGACCCCGCCAGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 293 GGTGTGGACCCCGCCAGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 352
QY 181 CTGCTCCAGCGCGGCTTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGAGCGGCCAGCA 240
Db 353 CTGCTCCAGCGCGGCTTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGAGCGGCCAGCA 412
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGCTTCCAGGCCACATCTCC 300
Db 413 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGCTTCCAGGCCACATCTCC 472
QY 301 CCCCGGCGCTGGCGAGGCCACCATCGAGTCCCAAGCCAGTGGGCCATCTCAGATGCGAG 360

Db 473 CCCGGGCTGGCGAGGCCACCATCGAGTCCACCACTGGCCATCTCAGATGCAGAG 532
QY 361 GACTGGTGCAGCTGAACAGATACAGCTGCAGAGTGCAGATGGCAGGGTGCCTACGGT 420
Db 533 GACTGGTGCAGCTGAACAGATACAGCTGCAGAGTGCAGATGGCAGGGTGCCTACGGT 592
QY 421 GTGGTGAAGGCTGACCTGACCAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCCCTTTCC 480
Db 593 GTGGTGAAGGCTGACCTGACCAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCCCTTTCC 652
QY 481 AAAAAGAAAGTACTGAAGCAAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGGTCCCGAG 540
Db 653 AAAAAGAAAGTACTGAAGCAAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGGTCCCGAG 712
QY 541 GCTGCCAGGAGGACCAAGCAGAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGAGATT 600
Db 713 GCTGCCAGGAGGACCAAGCAGAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGAGATT 772
QY 601 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGTCCTGGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGTCCTGGATGAC 832
QY 661 CCAGCTGAGACAACTCTATTTGGTGTGGTCTCTGAGAAAGGGGCCCGTCAATGGAA 720
Db 833 CCAGCTGAGACAACTCTATTTGGTGTGGTCTCTGAGAAAGGGGCCCGTCAATGGAA 892
QY 721 GTGCCCTGACAGAGCCCTTCTCGGAGGACAGCTGCGCTCTACCTCGCGGACGTCAATC 780
Db 893 GTGCCCTGACAGAGCCCTTCTCGGAGGACAGCTGCGCTCTACCTCGCGGACGTCAATC 952
QY 781 CTGGGCTCGAGTACTTGGCACTGCGAGAGATCTCCACAGGGACATCAAGCCATCCAAAC 840
Db 953 CTGGGCTCGAGTACTTGGCACTGCGAGAGATCTCCACAGGGACATCAAGCCATCCAAAC 1012
QY 841 CTGCTCTCGGGGATGATGGGACAGTGAAGATCGCCGACTTTGGCGTCAAGAACAGTTT 900
Db 1013 CTGCTCTCGGGGATGATGGGACAGTGAAGATCGCCGACTTTGGCGTCAAGAACAGTTT 1072
QY 901 GAGGGACACAGCTGAGCTGTCAGACAGCGGGGACCCAGCATTCATGGCCCCCGAG 960
Db 1073 GAGGGACACAGCTGAGCTGTCAGACAGCGGGGACCCAGCATTCATGGCCCCCGAG 1132
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGCTTTGGATGATGGGCCACTGGC 1020
Db 1133 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGCTTTGGATGATGGGCCACTGGC 1192
QY 1021 GTCAGTTGTAAGTCTGTTGCTATGGGAGTGGCCATTCATCGACGATTCATCTGGCC 1080
Db 1193 GTCAGTTGTAAGTCTGTTGCTATGGGAGTGGCCATTCATCGACGATTCATCTGGCC 1252
QY 1081 CTCACAGGAGATCAAGATGAGCCGCTGCTGTTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAGATCAAGATGAGCCGCTGCTGTTCTCTGAGGAGCCAGAAATCAGCGAG 1312
QY 1141 GAGCTCAAGGACCTGATCCTCAAGATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTG 1200
Db 1313 GAGCTCAAGGACCTGATCCTCAAGATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTG 1372
QY 1201 CCAGACATCAAGTGTGACCCCTTGGGTGACCAAGAACCGGGGAGGAGCCCTTCTTCGGAG 1260
Db 1373 CCAGACATCAAGTGTGACCCCTTGGGTGACCAAGAACCGGGGAGGAGCCCTTCTTCGGAG 1432
QY 1261 GAGGAGCACTCAGCGTGGTGGAGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTCAGCGTGGTGGAGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1492
QY 1321 ATCCCACTGACCAACAGCTGATCTCTGTGAGTCCATGCTGAGGAAGCGTTCTTTGGG 1380
Db 1493 ATCCCACTGACCAACAGCTGATCTCTGTGAGTCCATGCTGAGGAAGCGTTCTTTGGG 1552
QY 1381 AACCCGTTTGAAGCCCAAGGCAAGGAGGAAAGAGCGATCCATGTCTGCTCCAGGAAACCTA 1440
Db 1553 AACCCGTTTGAAGCCCAAGGCAAGGAGGAAAGAGCGATCCATGTCTGCTCCAGGAAACCTA 1612

QY 1441 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAA 1500
Db 1613 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687
RESULT 3
US-10-333-314-39
; Sequence 39, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURUPAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.; CHANLA, Narinder K.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: GANDHI, Ameena R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga
; APPLICANT: NGUYEN, Daniel B.; LU, Yan
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: DING, Li; YAO, Monique G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GLEZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 1937
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2190612CB1
US-10-333-314-39
Query Match 99.7%; Score 1510.2; DB 13; Length 1937;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGAGAGGGGTCCAGCTGTCTGCTGCCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 60
Db 65 ATGAGAGGGGTCCAGCTGTCTGCTGCCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 124
QY 61 GCACCCATCGATGACATCACTTGGAGGAGCAGATGTCGCCCCAGAGCCCTACTAGAAAC 120
Db 125 GCACCCATCGATGACATCACTTGGAGGAGCAGATGTCGCCCCAGAGCCCTACTAGAAAC 184
QY 121 GGTGTGACACCCCCACACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGACACCCCCACACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGCTCCAGCGCCCGCTAGCCTCTCAGCCAGGAAGCTTTCCTCTACAGGAGCGGCCAGCA 240

Db 245 CTGCTCCAGCCGCGCTAGCCCTCAGCCAGAGAGCTTCCCTACAGGAGGGCCAGCA 304
QY 241 GGAAGCTATCTGAGGGCGAGGTGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 300
Db 305 GGAAGCTATCTGAGGGCGAGGTGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 364
QY 301 CCCCGGCGCTGGGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCGAG 360
Db 365 CCCCGGCGCTGGGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCGAG 424
QY 361 GACTGCGTGCAGCTGAACCAAGTCAAGCTGCAGAGTCAGATTGGCAAGGGTGCTACGGT 420
Db 425 GACTGCGTGCAGCTGAACCAAGTCAAGCTGCAGAGTCAGATTGGCAAGGGTGCTACGGT 484
QY 421 GTGGTGAGGCTGCCCTACACGAAGTGAAGACAGACATATGCAATGAAGTCCCTTTC 480
Db 485 GTGGTGAGGCTGCCCTACACGAAGTGAAGACAGACATATGCAATGAAGTCCCTTTC 544
QY 481 AAAAAAGATTACTGAAGCAGTATGGCTTTCACAGTCGCCCTCCCGAGAGGGTCCAG 540
Db 545 AAAAAAGATTACTGAAGCAGTATGGCTTTCACAGTCGCCCTCCCGAGAGGGTCCAG 604
QY 541 GCTGCCAGGGAGCACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
Db 605 GCTGCCAGGGAGCACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 664
QY 601 GCATCTGTAAGAGCTGACACCGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 660
Db 665 GCATCTGTAAGAGCTGACACCGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 724
QY 661 CAGCTGAGGACAACTTATTTGGTGTGTGACTCTCGAAGAGGGGCCGCTCATGGA 720
Db 725 CAGCTGAGGACAACTTATTTGGTGTGTGACTCTCGAAGAGGGGCCGCTCATGGA 784
QY 721 GTGCCCTGTGAAGCCCTTCGGAGGACGAGCTCGCCTCTACCTGGGGACGCTCATC 780
Db 785 GTGCCCTGTGAAGCCCTTCGGAGGACGAGCTCGCCTCTACCTGGGGACGCTCATC 844
QY 781 CTGGGCGCTGAGTACTTGGACGTGCGAGATCGCCGACTTTGGCGTCAGCAACCACTT 840
Db 845 CTGGGCGCTGAGTACTTGGACGTGCGAGATCGCCGACTTTGGCGTCAGCAACCACTT 904
QY 841 CTGCTCCTGGGGGATGATGGGCACGTGAAGATCGCGACTTTGGCGTCAGCAACCACTT 900
Db 905 CTGCTCCTGGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCACTT 964
QY 901 GAGGGGAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 960
Db 965 GAGGGGAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 1024
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGATGGGCCACTGGC 1020
Db 1025 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGATGGGCCACTGGC 1084
QY 1021 GTCAAGTTGTACTGTCTTGTCTATGGGAAGTCCCATTCATCGAGATTTCATCTGGCC 1080
Db 1085 GTCAAGTTGTACTGTCTTGTCTATGGGAAGTCCCATTCATCGAGATTTCATCTGGCC 1144
QY 1081 CTCCACAGAGATCAAGATGAGCCCGTGTGTTCTCTGAGGACAGAAATCAGCGAG 1140
Db 1145 CTCCACAGAGATCAAGATGAGCCCGTGTGTTCTCTGAGGACAGAAATCAGCGAG 1204
QY 1141 GAGCTCAAGGACTCATCTGAAGATGTAGACAGAAATCCCGAGACGAGATTGGGGTG 1200
Db 1205 GAGCTCAAGGACTCATCTGAAGATGTAGACAGAAATCCCGAGACGAGATTGGGGTG 1264
QY 1201 CCAGACATCAAGTTGCACCTTGGGTGACCAAGAACGGGAGAGCCCTTCCCTCGGAG 1260
Db 1265 CCAGACATCAAGTTGCACCTTGGGTGACCAAGAACGGGAGAGCCCTTCCCTCGGAG 1324
QY 1261 GAGGAGCACTGACGCTGGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

Db 1325 GAGGAGCACTGCAGCTGGTGGAGGTGACAGAGAGAGGTTAAGAACTCAGTCAGGCTC 1384
QY 1321 ATCCCAGCTGGACCAACGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1380
Db 1385 ATCCCAGCTGGACCAACGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1444
QY 1381 AACCCGTTTGAAGCCCCAGGACCGAGGAAAGAGCGATCCATGTCTGCAGGAAACCTA 1440
Db 1445 AACCCGTTTGAAGCCCCAAGCACGAGGAAAGAGCGATCCATGTCTGCAGGAAACCTA 1504
QY 1441 CTGGTGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1500
Db 1505 CTGGTGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1564
QY 1501 GACGAGGCTGCATCC 1515
Db 1565 GACGAGGCTGCATCC 1579

RESULT 4

US-10-302-172-433
; Sequence 433, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCIP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 433
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1746)
US-10-302-172-433

Query Match 99.5%; Score 1507; DB 13; Length 2018;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 1510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGAGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGTGTAAGCGGGT 60
Db 232 ATGAGAGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGTGTAAGCGGGT 291
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTTACTAGAAC 120
Db 292 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTTACTAGAAC 351
QY 121 GGTGTGACCCCCCAACACCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 352 GGTGTGACCCCCCAACACCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 411
QY 181 CTGCTCCAGCCCGGCTTAGCTCTTCAGCCAGGAGCTTTCCTACAGAGCGGCCAGCA 240
Db 412 CTGCTCCAGCCCGGCTTAGCTCTTCAGCCAGGAGCTTTCCTACAGAGCGGCCAGCA 471
QY 241 GGAAGCTATCTGAGGCGCAGCTGGGCTTATGCGACGGGCGCTGCGAGCCACATCTCC 300
Db 472 GGAAGCTATCTGAGGCGCAGCTGGGCTTATGCGACGGGCGCTGCGAGCCACATCTCC 531

QY 301 CCCGGGCTGGCGAGGCCCAACATCGAGTCCACACAGTGGCCATCTCAGATGCAG 360
DB 532 CCCGGGCTGGCGAGGCCCAACATCGAGTCCACACAGTGGCCATCTCAGATGCAG 591
QY 361 GACTGCGTGCAGCTGAACACAGTCAAGTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 420
DB 592 GACTGCGTGCAGCTGAACACAGTCAAGTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 651
QY 421 GTGTGAGGCTGGCTCAACAGAAAGTGAACACAGACACTATGCAATGAAAGTCTTTCC 480
DB 652 GTGTGAGGCTGGCTCAACAGAAAGTGAACACAGACACTATGCAATGAAAGTCTTTCC 711
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGGAGAGGTTCCCG 540
DB 712 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGGAGAGGTTCCCG 771
QY 541 GCTGCCCAGGAGACAGCCCAAGCAGTGTGCCCCCTGGAGCGGGTGTACAGAGATT 600
DB 772 GCTGCCCAGGAGACAGCCCAAGCAGTGTGCCCCCTGGAGCGGGTGTACAGAGATT 831
QY 601 GCCATCCTGAAGAAAGCTGGACACGTCGAATGTGTCAAACCTGATCGAGTCTCGATGAC 660
DB 832 GCCATCCTGAAGAAAGCTGGACACGTCGAATGTGTCAAACCTGATCGAGTCTCGATGAC 891
QY 661 CCAGCTGAGGACAACTCTATTGTTGGTGTTCACCTCTCGAAGAGGGCCCGTCATGAA 720
DB 892 CCAGCTGAGGACAACTCTATTGTTGGTGTTCACCTCTCGAAGAGGGCCCGTCATGAA 951
QY 721 GTGCCCTGTACAAAGCCCTTCTCGAGGAGCAAGCTGCCCCCTACCTGCGGAGTCAATC 780
DB 952 GTGCCCTGTACAAAGCCCTTCTCGAGGAGCAAGCTGCCCCCTACCTGCGGAGTCAATC 1011
QY 781 CTGGCCCTCGAGTACTTGTCACTGCCAGAAAGTCTGCCACAGGACATCAAGCCATCAAC 840
DB 1012 CTGGCCCTCGAGTACTTGTCACTGCCAGAAAGTCTGCCACAGGACATCAAGCCATCAAC 1071
QY 841 CTGTCTCTGGGGATGATGGGCACTGGAAGATCGCCGACTTTGGCGTGAAGCAACAGTTT 900
DB 1072 CTGTCTCTGGGGATGATGGGCACTGGAAGATCGCCGACTTTGGCGTGAAGCAACAGTTT 1131
QY 901 GAGGGAAACGAGCTGAGCTGTCCAGCAGCGCGGAACCCAGCATTCATGGCCCCCGAG 960
DB 1132 GAGGGAAACGAGCTGAGCTGTCCAGCAGCGCGGAACCCAGCATTCATGGCCCCCGAG 1191
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTTGGATGTATGGGCCACTGGC 1020
DB 1192 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTTGGATGTATGGGCCACTGGC 1251
QY 1021 GTCAAGTTGTACTGTTTGTATGGGAAGTCCCATTCATCGACGATTCATCTGGCC 1080
DB 1252 GTCAAGTTGTACTGTTTGTATGGGAAGTCCCATTCATCGACGATTCATCTGGCC 1311
QY 1081 CTCACAGGAGATCAAGATGAGCCGTGTGTTTCTGAGGAGCAGAAATCAGCAG 1140
DB 1312 CTCATAGGAAGATCAAGATGAGCCGTGTGTTTCTGAGGAGCAGAAATCAGCAG 1371
QY 1141 GAGCTCAAGGACCTGATCTTGAAGATGTAGACAAGAAATCCGAGACAGAAATGGGGT 1200
DB 1372 GAGCTCAAGGACCTGATCTTGAAGATGTAGACAAGAAATCCGAGACAGAAATGGGGT 1431
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCCTTCGGAG 1260
DB 1432 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCCTTCGGAG 1491
QY 1261 GAGGACACTGCAGCGTGTGAGGTGACAGAGGGGAGGTTAGAACTAGTCAGCTC 1320
DB 1492 GAGGACACTGCAGCGTGTGAGGTGACAGAGGGGAGGTTAGAACTAGTCAGCTC 1551
QY 1321 ATCCCCAGCTGGACACAGCTGATCTGTGTAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1380
DB 1552 ATCCCCAGCTGGACACAGCTGATCTGTGTAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1611
QY 1381 AACCCGTTTGAGCCCCAGGCAAGGAGGAGGAGCGATCCATGCTGCTCCAGGAACCTA 1440

DB 1612 AACCCGTTTGAGCCCCAAGCACCGGAGGAGGATCCATGCTGCTCCAGGAACCTA 1671
QY 1441 CTGSGTGAAGAGGGTTTGGTGAAGGGGGGCAAGAGCCAGAGCTCCCGCGGCTCCAGGAA 1500
DB 1672 CTGSGTGAAGAGGGTTTGGTGAAGGGGGGCAAGAGCCAGAGCTCCCGCGGCTCCAGGAA 1731
QY 1501 GACGAGGCTGCATCC 1515
DB 1732 GACGAGGCTGCATCC 1746

RESULT 5
US-10-092-900A-55
; Sequence 55, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Peña, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.

Qy	1	ATGAGGGGGTTCAGCTGTCTGCTGCCAGGATCTCTGGG	CAGAGCTGTGGT	AAGACCGGTG	60
Db	1	ATGAGGGGGTTCAGCTGTCTGCTGCCAGGATCTCTGGG	CAGAGCTGTGGT	AAGACCGGTG	60
Qy	61	GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGGT	GGCCAGAGCCTACT	TAGAAC	120
Db	61	GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGGT	GGCCAGAGCCTACT	TAGAAC	120
Qy	121	GGTGTGACCCCCACCA	CGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA	180	
Db	121	GGTGTGACCCCCACCA	CGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA	180	
Qy	181	CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGAGAGCTTCCCTACAGGAGCGGCAGCA	240		
Db	181	CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGAGAGCTTCCCTACAGGAGCGGCAGCA	240		
Qy	241	GGAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCACGGGGCTGCCAGCCATCTCC	300		
Db	241	GGAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCACGGGGCTGCCAGCCATCTCC	300		
Qy	301	CCC	CGGGCTGGCGGAGGCCACCATCGAGTCCCA	CCACGTGGCCATCTCAGATGCGAG	360
Db	301	CCC	CGGGCTGGCGGAGGCCACCATCGAGTCCCA	CCACGTGGCCATCTCAGATGCGAG	360
Qy	361	GACTGTGCTGAGTGAACCACTA	CAAGTGCAGAGTGAGATTGGCAAGGTGCCCTACGGT	420	
Db	361	GACTGTGCTGAGTGAACCACTA	CAAGTGCAGAGTGAGATTGGCAAGGTGCCCTACGGT	420	
Qy	421	GTGTGAGGCTGGCTCAACGAAAGTGAA	CACAGACACTATGCATGAAAGTCTCTTCC	480	
Db	421	GTGTGAGGCTGGCTCAACGAAAGTGAA	CACAGACACTATGCATGAAAGTCTCTTCC	480	
Qy	481	AAAAAGAGTTACTCAAGCAGTATGGCTTCC	ACAGTCCGCCCTCCCGAGAGGTCCCGAG	540	
Db	481	AAAAAGAGTTACTCAAGCAGTATGGCTTCC	ACAGTCCGCCCTCCCGAGAGGTCCCGAG	540	
Qy	541	GCTGCCCAGGAGACCA	CGCAAGCAGTGTGCTGCCCTGGAGCGGTGTATC	CAGAGATT	600
Db	541	GCTGCCCAGGAGACCA	CGCAAGCAGTGTGCTGCCCTGGAGCGGTGTATC	CAGAGATT	600
Qy	601	GCCATCTGAAGAAGCTTGGACCA	CGTGAATGTGTCAAACTGATCGAGGTCTCGATGAC	660	
Db	601	GCCATCTGAAGAAGCTTGGACCA	CGTGAATGTGTCAAACTGATCGAGGTCTCGATGAC	660	
Qy	661	CCAGCTCAGGACAACTCTATTGGTGT	TTCACCTCTGCAAAAGGGCCCGTCA	TGAA	720
Db	661	CCAGCTCAGGACAACTCTATTGGTGT	TTCACCTCTGCAAAAGGGCCCGTCA	TGAA	720
Qy	721	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCTCTACCTCGGGAGCTCATC	780		
Db	721	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCTCTACCTCGGGAGCTCATC	780		
Qy	781	CTGGGCTCGAGTACTTGCATCGCAGAGATCGTCCA	CAGGACATCAAGCCATCCAAC	840	
Db	781	CTGGGCTCGAGTACTTGCATCGCAGAGATCGTCCA	CAGGACATCAAGCCATCCAAC	840	
Qy	841	CTGCTCTCGGGGATGATGGGCAGTGAAGATCGCGA	CTTTGGCGTCAGCAACCGATT	900	
Db	841	CTGCTCTCGGGGATGATGGGCAGTGAAGATCGCGA	CTTTGGCGTCAGCAACCGATT	900	
Qy	901	GAGGGAA	CGACGTCTCAGCA	CGCGGGAACCCAGCATTCATGCGCCCCGAG	960
Db	901	GAGGGAA	CGACGTCTCAGCA	CGCGGGAACCCAGCATTCATGCGCCCCGAG	960
Qy	961	GCAATTTCTGATT	CGGGCAGACTT	CAGTGGGAAGGCTTGGATGTATGGGCCACTGGC	1020
Db	961	GCAATTTCTGATT	CGGGCAGACTT	CAGTGGGAAGGCTTGGATGTATGGGCCACTGGC	1020
Qy	1021	GTCAGTTGTACTCTTTGCTAT	TGGGAAGTGCCCATTCATCGACGATTTT	CATCTCGCC	1080
Db	1021	GTCAGTTGTACTCTTTGCTAT	TGGGAAGTGCCCGTTTCATCGACGATTTT	CATCTCGCC	1080
Qy	1081	CTCCA	CAGGAAGATCAAGAATGAGCCCTGTGTTTCTCGAGGAGCGAGAAATCAGCGAG	1140	

RESULT 7

RESULTS / IIS-10-391-364-43

```

; Sequence 43, Application US/10391364
; Publication No. US20040121349A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Weich, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14000, 14001, 14002, 14003, 14004, 14005, 14006, 14007, 14008, 14009, 14010, 14011, 14012, 14013, 14014, 14015, 14016, 14017, 14018, 14019, 14020, 14021, 14022, 14023, 14024, 14025, 14026, 14027, 14028, 14029, 14030, 14031, 14032, 14033, 14034, 14035, 14036, 14037, 14038, 14039, 14040, 14041, 14042, 14043, 14044, 14045, 14046, 14047, 14048, 14049, 14050, 14051, 14052, 14053, 14054, 14055, 14056, 14057, 14058, 14059, 14060, 14061, 14062, 14063, 14064, 14065, 14066, 14067, 14068, 14069, 14070, 14071, 14072, 14073, 14074, 14075, 14076, 14077, 14078, 14079, 14080, 14081, 14082, 14083, 14084, 14085, 14086, 14087, 14088, 14089, 14090, 14091, 14092, 14093, 14094, 14095, 14096, 14097, 14098, 14099, 14100, 14101, 14102, 14103, 14104, 14105, 14106, 14107, 14108, 14109, 14110, 14111, 14112, 14113, 14114, 14115, 14116, 14117, 14118, 14119, 14120, 14121, 14122, 14123, 14124, 14125, 14126, 14127, 14128, 14129, 14130, 14131, 14132, 14133, 14134, 14135, 14136, 14137, 14138, 14139, 14140, 14141, 14142, 14143, 14144, 14145, 14146, 14147, 14148, 14149, 14150, 14151, 14152, 14153, 14154, 14155, 14156, 14157, 14158, 14159, 14160, 14161, 14162, 14163, 14164, 14165, 14166, 14167, 14168, 14169, 14170, 14171, 14172, 14173, 14174, 14175, 14176, 14177, 14178, 14179, 14180, 14181, 14182, 14183, 14184, 14185, 14186, 14187, 14188, 14189, 14190, 14191, 14192, 14193, 14194, 14195, 14196, 14197, 14198, 14199, 14200, 14201, 14202, 14203, 14204, 14205, 14206, 14207, 14208, 14209, 14210, 14211, 14212, 14213, 14214, 14215, 14216, 14217, 14218, 14219, 14220, 14221, 14222, 14223, 14224, 14225, 14226, 14227, 14228, 14229, 14230, 14231, 14232, 14233, 14234, 14235, 14236, 14237, 14238, 14239, 14240, 14241, 14242, 14243, 14244, 14245, 14246, 14247, 14248, 14249, 14250, 14251, 14252, 14253, 14254, 14255, 14256, 14257, 14258, 14259, 14260, 14261, 14262, 14263, 14264, 14265, 14266, 14267, 14268, 14269, 14270, 14271, 14272, 14273, 14274, 14275, 14276, 14277, 14278, 14279, 14280, 14281, 14282, 14283, 14284, 14285, 14286, 14287, 14288, 14289, 14290, 14291, 14292, 14293, 14294, 14295, 14296, 14297, 14298, 14299, 14300, 14301, 14302, 14303, 14304, 14305, 14306, 14307, 14308, 14309, 14310, 14311, 14312, 14313, 14314, 14315, 14316, 14317, 14318, 14319, 14320, 14321, 14322, 14323, 14324, 14325, 14326, 14327, 14328, 14329, 14330, 14331, 14332, 14333, 14334, 14335, 14336, 14337, 14338, 14339, 14340, 14341, 14342, 14343, 14344, 14345, 14346, 14347, 14348, 14349, 14350, 14351, 14352, 14353, 14354, 14355, 14356, 14357, 14358, 14359, 14360, 14361, 14362, 14363, 14364, 14365, 14366, 14367, 14368, 14369, 14370, 14371, 14372, 14373, 14374, 14375, 14376, 14377, 14378, 14379, 14380, 14381, 14382, 14383, 14384, 14385, 14386, 14387, 14388, 14389, 14390, 14391, 14392, 14393, 14394, 14395, 14396, 14397, 14398, 14399, 14400, 14401, 14402, 14403, 14404, 14405, 14406, 14407, 14408, 14409, 14410, 14411, 14412, 14413, 14414, 14415, 14416, 14417, 14418, 14419, 14420, 14421, 14422, 14423, 14424, 14425, 14426, 14427, 14428, 14429, 14430, 14431, 14432, 14433, 14434, 14435, 14436, 14437, 14438, 14439, 14440, 14441, 14442, 14443, 14444, 14445, 14446, 14447, 14448, 14449, 14450, 14451, 14452, 14453, 14454, 14455, 14456, 14457, 14458, 14459, 14460, 14461, 14462, 14463, 14464, 14465, 14466, 14467, 14468, 14469, 14470, 14471, 14472, 14473, 14474, 14475, 14476, 14477, 14478, 14479, 14480, 14481, 14482, 14483, 14484, 14485, 14486, 14487, 14488, 14489, 14490, 14491, 14492, 14493, 14494, 14495, 14496, 14497, 14498, 14499, 14500, 14501, 14502, 14503, 14504, 14505, 14506, 14507, 14508, 14509, 14510, 14511, 14512, 14513, 14514, 14515, 14516, 14517, 14518, 14519, 14520, 14521, 14522, 14523, 14524, 14525, 14526, 14527, 14528, 14529, 14530, 14531, 14532, 14533, 14534, 14535, 14536, 14537, 14538, 14539, 14540, 14541, 14542, 14543, 14544, 14545, 14546, 14547, 14548, 14549, 14550, 14551, 14552, 14553, 14554, 14555, 14556, 14557, 14558, 14559, 14560, 14561, 14562, 14
```

```
; SEQ ID NO 43
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1686)
US-10-391-364-43

Query Match      98.4%; Score 1490.6; DB 17; Length 1686;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCGAGAGCTGGTAGAACGGGTG 60
Db 1 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCGAGAGCTGGTAGAACGGGTG 60

QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGCAGATGGTGGCCCGAGAGCCTACTAGAAAC 120
Db 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGCAGATGGTGGCCCGAGAGCCTACTAGAAAC 120

QY 121 GGTGTGGACCCCCACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCGAGTACTCAAGA 180
Db 121 GGTGTGGACCCCCACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCGAGTACTCAAGA 180

QY 181 CTGCTCCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db 181 CTGCTCCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240

QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTCCCGACCATCTCC 300
Db 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTCCCGACCATCTCC 300

QY 301 CCCCGGGCTGGGGAGGCCACCATCGAGTCCCAACCAAGTGGCCATCTCAGATGCAGAG 360
Db 301 CCCCGGGCTGGGGAGGCCACCATCGAGTCCCAACCAAGTGGCCATCTCAGATGCAGAG 360

QY 361 GACTGGGTGAGCTGAACAGTACAACTGACAGTGCAGATGGCAAGGGTGCCTACGGT 420
Db 361 GACTGGGTGAGCTGAACAGTACAACTGACAGTGCAGATGGCAAGGGTGCCTACGGT 420

QY 421 GTGGTGAAGCTGGCTTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480
Db 421 GTGGTGAAGCTGGCTTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480

QY 481 AAAAAAGTGTACTGAAGCAGTATGGCTTTCCAGTCCGCCCTCCCGAGAGGGTCCCGAG 540
Db 481 AAAAAAGTGTACTGAAGCAGTATGGCTTTCCAGTCCGCCCTCCCGAGAGGGTCCCGAG 540

QY 541 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600
Db 541 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600

QY 601 GCATCTCTGAAGAAGCTGACACAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 660
Db 601 GCATCTCTGAAGAAGCTGACACAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 660

QY 661 CCAGCTGAGGAGCAACTCTATTGTTGGTTGACCTCTCAGAAAGGGGCCCTCATGGAA 720
Db 661 CCAGCTGAGGAGCAACTCTATTGTTGGTTGACCTCTCAGAAAGGGGCCCTCATGGAA 720

QY 721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGGAGCGTCATC 780
Db 721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGGAGCGTCATC 780

QY 781 CTGGGCTCGAGTACTGTCAGTCCAGAGATCGTCCACAGGAGATCAAGCATCCAC 840
Db 781 CTGGGCTCGAGTACTGTCAGTCCAGAGATCGTCCACAGGAGATCAAGCATCCAC 840

QY 841 CTGCTCCTGGGGATGATGGGACGTGAAGATCGCCGACTTTGGCGTTCAGACCAAGTTT 900
Db 841 CTGCTCCTGGGGATGATGGGACGTGAAGATCGCCGACTTTGGCGTTCAGACCAAGTTT 900
```

```
QY 901 GAGGGGAACGACGCTCAGCTGTCCAGACGGGGGACCCACAGATTTCATGGCCCCCGAG 960
Db 901 GAGGGGAACGACGCTCAGCTGTCCAGACGGGGGACCCACAGATTTCATGGCCCCCGAG 960

QY 961 GCCATTTCGATTCCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCACTGGC 1020
Db 961 GCCATTTCGATTCCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCACTGGC 1020

QY 1021 GTCACGTTGCTACTGCTTTTCTATGGGAAGTGCCTATTCATCGACGATTTCATCTGGCC 1080
Db 1021 GTCACGTTGCTACTGCTTTTCTATGGGAAGTGCCTATTCATCGACGATTTCATCTGGCC 1080

QY 1081 CTCACAGGAAGATCAAGATGAGCCCGTGGTTCCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1081 CTCACAGGAAGATCAAGATGAGCCCGTGGTTCCTCTGAGGAGCCAGAAATCAGCGAG 1140

QY 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATGGGGTG 1200
Db 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATGGGGTG 1200

QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCGGAG 1260
Db 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCGGAG 1260

QY 1261 GAGGACACTGACAGCGTGGTGGAGTGCAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1261 GAGGACACTGACAGCGTGGTGGAGTGCAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

QY 1321 ATCCCCAGCTGGAACCAACGCTGATCCTGTGAAGTCCATCTGAGGAAGGCTTCTTTGGG 1380
Db 1321 ATCCCCAGCTGGAACCAACGCTGATCCTGTGAAGTCCATCTGAGGAAGGCTTCTTTGGG 1380

QY 1381 AACCGTTTGGAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381 AACCGTTTGGAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

QY 1441 CTGGTGAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCCGGCGTCCAGGAA 1500
Db 1441 CTGGTGAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCCGGCGTCCAGGAA 1500

QY 1501 GACGA 1505
Db 1501 TACCA 1505

RESULT 8
US-09-922-138-7
; Sequence 7, Application US/09922138
; Patent No. US20020061574A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 16658, 14223, AND 16002, NOVEL HUMAN
; TITLE OF INVENTION: KINASES AND USES THEREFOR
; FILE REFERENCE: 38155-20030.00
; CURRENT APPLICATION NUMBER: US/09/922,138
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/229,299
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (198)...(1880)
; NAME/KEY: misc feature
; LOCATION: (1)-(2711)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-138-7
```

Query Match 98.4%; Score 1490.6; DB 9; Length 2711;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ATGAGGGGGTCCAGTGTCTGTCCAGGATCTCTGGGAGAGCTGGTAGAACGGGTG 60
Db ATGAGGGGGTCCAGTGTCTGTCCAGGATCTCTGGGAGAGCTGGTAGAACGGGTG 257

Qy 61 GCAGCCATCATGTGACTCACTTGGAGAGGCAGATGTCGCCAGAGCCCTACTAGAAC 120
Db GCAGCCATCATGTGACTCACTTGGAGAGGCAGATGTCGCCAGAGCCCTACTAGAAC 317

Qy 121 GGTGTGACCCCCACCACCGGGCCAGAGCTGCTGTGTATCCCTGGCAGTACTCAAGA 180
Db GGTGTGACCCCCACCACCGGGCCAGAGCTGCTGTGTATCCCTGGCAGTACTCAAGA 377

Qy 181 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
Db CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 437

Qy 241 GGAAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCAGGGGCTGCCAGCATCTCC 300
Db GGAAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCAGGGGCTGCCAGCATCTCC 497

Qy 301 CCCGGGCTGGGGAGGCGCCACCATCGAGTCCACCAACCGTGGCCATCTCAGATGCAGAG 360
Db CCCGGGCTGGGGAGGCGCCACCATCGAGTCCACCAACCGTGGCCATCTCAGATGCAGAG 557

Qy 361 GACTGCTGCAGCTGAACCAAGTCAAGCTGCAGAGTGGAGTGGCAAGGGTGCCTACGGT 420
Db GACTGCTGCAGCTGAACCAAGTCAAGCTGCAGAGTGGAGTGGCAAGGGTGCCTACGGT 617

Qy 421 GTGTGAGGCTGGCTACACGAAGTGAAGACAGACTATGCAATGAAGTCTCTTCC 480
Db GTGTGAGGCTGGCTACACGAAGTGAAGACAGACTATGCAATGAAGTCTCTTCC 677

Qy 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCCGAG 540
Db AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCCGAG 737

Qy 541 GCTGCCAGGAGGACGAGCCAGCAGCTGTGCTGCTGGAGCGGGTGTACAGAGATT 600
Db GCTGCCAGGAGGACGAGCCAGCAGCTGTGCTGCTGGAGCGGGTGTACAGAGATT 797

Qy 601 GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACTGATCGAGGCTCTGGATGAC 660
Db GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACTGATCGAGGCTCTGGATGAC 857

Qy 661 CCAGCTGAGGACAACTCTATTGTTGGTGTGACCTCCTCGAAGAGGGGCGCGTCATGAA 720
Db CCAGCTGAGGACAACTCTATTGTTGGTGTGACCTCCTCGAAGAGGGGCGCGTCATGAA 917

Qy 721 GTGGCTGTGACAGCCCTTCTCGAGGAGCAAGCTGCGCTCTACCTGCGGAGCTCATC 780
Db GTGGCTGTGACAGCCCTTCTCGAGGAGCAAGCTGCGCTCTACCTGCGGAGCTCATC 977

Qy 781 CTGGGCTCTGAGTACTTGTGCTGCAAGAGATGCTCCAGGGACATCAAGCCATCAAC 840
Db CTGGGCTCTGAGTACTTGTGCTGCAAGAGATGCTCCAGGGACATCAAGCCATCAAC 1037

Qy 841 CTGCTCTGGGGATGATGGGACAGTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 900
Db CTGCTCTGGGGATGATGGGACAGTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 1097

Qy 901 GAGGGGAACGACGCTCAGCTGTCCAGCAACGGCGGGAACCCAGCATTCATGCGCCCGAG 960
Db GAGGGGAACGACGCTCAGCTGTCCAGCAACGGCGGGAACCCAGCATTCATGCGCCCGAG 1157

Qy 961 GCCATTTCTGATTCGGGCCAGCTTCAAGTGGGAGGCTTGAATGTATGGCCACTGGC 1020
Db GCCATTTCTGATTCGGGCCAGCTTCAAGTGGGAGGCTTGAATGTATGGCCACTGGC 1217

RESULT 9

US-10-391-364-41
; Sequence 41, Application US/10391364
; Publication No. US20040121349A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: USFS THEREFOR
; FILE REFERENCE: MEI03-0190NMIM
; CURRENT APPLICATION NUMBER: US/10391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/266,036
; PRIOR FILING DATE: 2002-10-09

Qy 1021 GTCACGTTGATCTGCTTTGTCTATGGAGTGGCCATTCATCGACGATTTTCATCTCGCC 1080
Db GTCACGTTGATCTGCTTTGTCTATGGAGTGGCCGCTTCATCGACGATTTTCATCTCGCC 1277

Qy 1081 CTCACAGGAAGATCAAGATGAGCCCGTGGTGTCTTCTGAGGAGCCAGAAAATCAGCGAG 1140
Db CTCACAGGAAGATCAAGATGAGCCCGTGGTGTCTTCTGAGGAGCCAGAAAATCAGCGAG 1337

Qy 1141 GAGCTCAAGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTTGGGGTG 1200
Db GAGCTCAAGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTTGGGGTG 1397

Qy 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTTCGGAG 1260
Db CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTTCGGAG 1457

Qy 1261 GAGGAGCACTGCAGCGTGGTGGAGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db GAGGAGCACTGCAGCGTGGTGGAGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1517

Qy 1321 ATCCCCAGCTGGACCAACCGTGTATCTGCTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGG 1380
Db ATCCCCAGCTGGACCAACCGTGTATCTGCTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGG 1577

Qy 1381 AACCCGTTTTCAGCCCCAGGCAACGAGGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1440
Db AACCCGTTTTCAGCCCCAGGCAACGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1637

Qy 1441 CTGGTGAAGAGAGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGGGGTCAGAA 1500
Db CTGGTGAAGAGAGGTTTGGTGAAGGGGGCAAGAGCTCCCGGGGTCAGAGCT 1697

Qy 1501 GACGA 1505
Db 1598 TACCA 1702

```
/ PRIOR APPLICATION NUMBER: US 60/327,820
/ PRIOR FILING DATE: 2001-10-09
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 93
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41
/ LENGTH: 2711
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (198)...(1883)
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(2711)
/ OTHER INFORMATION: n = A,T,C or G
US-10-391-364-41

Query Match      98.4%; Score 1490.6; DB 17; Length 2711;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 ATGGAGGGGCTCCAGCTGTCTCTGCGAGGATCTCTCGGGCAGAGCTGGTAGAACGGGTG 60
DB      198 ATGGAGGGGGTCCAGCTGTCTCTGCGCAGGATCTCTCGGGCAGAGCTGGTAGAACGGGTG 257
QY      61 GCAGCCATCGATGTGACTCACTTGGAGGAGGCGAGATGGTGGCCACAGAGCTACTAGAAAC 120
DB      258 GCAGCCATCGATGTGACTCACTTGGAGGAGGCGAGATGGTGGCCACAGAGCTACTAGAAAC 317
QY      121 GGTGTGGAACCCGCCACCAAGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB      318 GGTGTGGAACCCGCCACCAAGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 377
QY      181 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGAGCGGCCAGCA 240
DB      378 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGAGCGGCCAGCA 437
QY      241 GGAAGTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTGCCAGCCACATCTCC 300
DB      438 GGAAGTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTGCCAGCCACATCTCC 497
QY      301 CCCGGGCTGGCGAGGCCACCATCGAGTCCCACCGTGGCCATCTCAGATCGAG 360
DB      498 CCCGGGCTGGCGAGGCCACCATCGAGTCCCACCGTGGCCATCTCAGATCGAG 557
QY      361 GACTCGTGCAGTGAACAGTACAGCTGCAGTGCAGTGGCAGAGGTCCTACGTT 420
DB      558 GACTCGTGCAGTGAACAGTACAGCTGCAGTGGCAGTGGCAGAGGTCCTACGTT 617
QY      421 GTGGTGAGGCTGGCCTTACCAAGAAAGTGAAGACAGACATATGCAATGAAAGTCTCTTC 480
DB      618 GTGGTGAGGCTGGCCTTACCAAGAAAGTGAAGACAGACATATGCAATGAAAGTCTCTTC 677
QY      481 AAAAAGATTACTGAAGCAGTATGGCTTTCCAGCTGGCCCTCCCGGAGAGGGTCCGAG 540
DB      678 AAAAAGAAATTACTGAAGCAGTATGGCTTTCCAGCTGGCCCTCCCGGAGAGGGTCCGAG 737
QY      541 GCTGCCAGGAGGACCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
DB      738 GCTGCCAGGAGGACCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 797
QY      601 GCCATCTTGAAGAAGCTGGACACCGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660
DB      798 GCCATCTTGAAGAAGCTGGACACCGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 857
QY      661 CCAGCTGAGGACAACTCTATTGTGTTTGTGCTTCTGAGAAAGGGGCCCTCATGGA 720
DB      858 CCAGCTGAGGACAACTCTATTGTGTTTGTGCTTCTGAGAAAGGGGCCCTCATGGA 917
QY      721 GTGGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCTCTACTCGGGACGCTCATC 780
DB      918 GTGGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCTCTACTCGGGACGCTCATC 977

QY      781 CTGGGCTCGACTACTTGCACCTGCCAGAGATCGTCCACAGGACATCAAGCCATCAAC 840
DB      978 CTGGGCTCGACTACTTGCACCTGCCAGAGATCGTCCACAGGACATCAAGCCATCAAC 1037
QY      841 CTGCTCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCACTTT 900
DB      1038 CTGCTCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCACTTT 1097
QY      901 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGAAACCCAGCATTTATGGCCCCGAG 960
DB      1098 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGAAACCCAGCATTTATGGCCCCGAG 1157
QY      961 GCCATTTCTGATTCCGGCCAGAGCTTCACTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1020
DB      1158 GCCATTTCTGATTCCGGCCAGAGCTTCACTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1217
QY      1021 GTACGTTGTACTGCTTTCTATGGGAAGTGCCTTTCATCGACGATTTTCATCTCTGGCC 1080
DB      1218 GTACGTTGTACTGCTTTCTATGGGAAGTGCCTTTCATCGACGATTTTCATCTCTGGCC 1277
QY      1081 CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
DB      1278 CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1337
QY      1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAAATTTGGGTG 1200
DB      1338 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAAATTTGGGTG 1397
QY      1201 CCAGACATCAAGTTGCACCTTTGGGTGACCAAGAAAGGGGAGGAGCCCTTCTCTCGGAG 1260
DB      1398 CCAGACATCAAGTTGCACCTTTGGGTGACCAAGAAAGGGGAGGAGCCCTTCTCTCGGAG 1457
QY      1261 GAGGAGCATCTGACGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB      1458 GAGGAGCATCTGACGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1517
QY      1321 ATCCCCAGCTGGACCAACCGGTGATCTGTTGAAAGTCCATCTCAGGAAGCGTTCCTTTGG 1380
DB      1518 ATCCCCAGCTGGACCAACCGGTGATCTGTTGAAAGTCCATCTCAGGAAGCGTTCCTTTGG 1577
QY      1381 AACCCGTTTACCCCCAGGCACGAGGAGGAGCGATCATGTCTCCAGAAACCTA 1440
DB      1578 AACCCGTTTACCCCCAGGCACGAGGAGGAGCGATCATGTCTCCAGAAACCTA 1637
QY      1441 CTGGTGAAGAGGGTGGTGAAGGGGCGAGAGCCAGAGCTCCCCGGGTCAGGAA 1500
DB      1638 CTGGTGAAGAGGGTGGTGAAGGGGCGAGAGCCAGAGCTCCCCGGGTCAGGCT 1697
QY      1501 GACGA 1505
DB      1698 TACCA 1702

RESULT 10
US-10-168-582-14
; Sequence 14, Application US/10168582
; Publication No. US20040058426A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa
; APPLICANT: YUE, Henry
; APPLICANT: YAO, Monique G.
; APPLICANT: LAL, Preeti
; APPLICANT: KAHN, Farrah A.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0002 PCT
; CURRENT APPLICATION NUMBER: US/10/168,582
```



```
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 60/172,066; 60/176,107; 60/176,107; 60/177,731
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-14; 2000-01-21
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058426A1 2041716CB1
US-10-168-582-14

Query Match          96.2%; Score 1457.2; DB 13; Length 3501;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 3; Indels 30; Gaps 2;

QY 1 ATGAGGGGGTCCAGTGTCTGTGTCAGGATCCTCGGGCAGAGCTGTGTAGAACCGGTG 60
DB 1 ATGAGGGGGTCCAGTGTCTGTGTCAGGATCCTCGGGCAGAGCTGTGTAGAACCGGTG 124
QY 61 GCACCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGGCCAGAGCCTACTAGAAC 120
DB 125 GCACCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGGCCAGAGCCTACTAGAAC 184
QY 121 GGTGTGACCCGCCACCAACGAGGCGCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 180
DB 185 GGTGTGACCCGCCACCAACGAGGCGCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGCTCCAGCGCGCTAGCCTCTACGACAGAGAACTTTCCTTACAGGAGCGGCCAGCA 240
DB 245 CTGCTCCAGCGCGCTAGCCTCTACGACAGAGAACTTTCCTTACAGGAGCGGCCAGCA 304
QY 241 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTTATGCCAGGGGCTGCCAGCCATCTCC 300
DB 305 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTTATGCCAGGGGCTGCCAGCCATCTCC 364
QY 301 CCCCGGCTCTGGGAGGCGCCACCATCGAGTCCACCAACGAGTGGCCATCTCAGATGCGAG 360
DB 365 CCCCGGCTCTGGGAGGCGCCACCATCGAGTCCACCAACGAGTGGCCATCTCAGATGCGAG 424
QY 361 GACTGCGTGCAGTGAACAGTACAGCTCAAGCTGAGAGTGAAGTGGCA----- 406
DB 425 GACTGCGTGCAGTGAACAGTACAGCTCAAGCTGAGAGTGAAGTGGCA----- 484
QY 407 -----AGGCTGCTACGCTGCTGGTGGAGGCTGGCTTACAGCAAGTGAAGTGAAG 453
DB 485 GATGCTTATCTGAGGGTGCTTACGCTGCTGGTGGAGGCTGGCTTACAGCAAGTGAAGTGAAG 544
QY 454 AGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTTACTGAAAGCAGTATGGCTTTCCA 513
DB 545 AGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTTACTGAAAGCAGTATGGCTTTCCA 604
QY 514 CGTGGCTCTCCCGAGAGGGTCCAGGCTGCGCCAGGAGGACAGCAGCAGAGCTGCTG 573
DB 605 CGTGGCTCTCCCGAGAGGGTCCAGGCTGCGCCAGGAGGACAGCAGCAGAGCTGCTG 664
QY 574 CCCTGGAGCGGGTGTACAGGAGATGGCCATCTGAGAGCTGGACCCAGCTGAATGTG 633
DB 665 CCCTGGAGCGGGTGTACAGGAGATGGCCATCTGAGAGCTGGACCCAGCTGAATGTG 724
QY 634 GTCAACTGATCGAGGCTCTGGATGACCCAGCTGAGGACCACTCTATTTTGGTGTGTGAC 693
DB 725 GTCAACTGATCGAGGCTCTGGATGACCCAGCTGAGGACCACTCTATTTTGG---TTGAC 781
QY 694 CTCCTGAGAAAGGGGCCCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGGAGAGCAA 753
DB 782 CTCCTGAGAAAGGGGCCCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGGAGAGCAA 841
QY 754 GCTCGCTCTACCTGCGGGAGCTCATCTGGGCTCGAGTACTTGCATCTGCCAGAGATC 813
DB 842 GCTCGCTCTACCTGCGGGAGCTCATCTGGGCTCGAGTACTTGCATCTGCCAGAGATC 901
```

```
QY 814 GTCACAGGACATCAAGCCATCAACCTGCTCTCTGGGGATGATGGCAGTGAAGATC 873
DB 902 GTCACAGGACATCAAGCCATCAACCTGCTCTCTGGGGATGATGGCAGTGAAGATC 961
QY 874 GCGACTTTGGCGTCAGCAACCAAGTTTGGGGGAAACGACGCTCAGCTGTCTCCAGCACGGCG 933
DB 962 GCGACTTTGGCGTCAGCAACCAAGTTTGGGGGAAACGACGCTCAGCTGTCTCCAGCACGGCG 1021
QY 934 GGAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTTCCGGGCCAGAGCTTCACTGGG 993
DB 1022 GGAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTTCCGGGCCAGAGCTTCACTGGG 1081
QY 994 AAGCCCTTGTATGTATGGGCCATCTGGCGTCACTGCTTGTACTGCTTGTCTATGGGAAGTGC 1053
DB 1082 AAGCCCTTGTATGTATGGGCCATCTGGCGTCACTGCTTGTACTGCTTGTCTATGGGAAGTGC 1141
QY 1054 CCATTATCAACGATTTTCATCTTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTG 1113
DB 1142 CCGTTTCATCGAGATTTTCATCTTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTG 1201
QY 1114 TTTCTCAGGAGCGCAGAAATCAGCGAGAGCTCAAGAACCTGATCTCTGAAGATGTTAGAC 1173
DB 1202 TTTCTCAGGAGCGCAGAAATCAGCGAGAGCTCAAGAACCTGATCTCTGAAGATGTTAGAC 1261
QY 1174 AAGAATCCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGCACTTGGGTGACCAAG 1233
DB 1262 AAGAATCCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGCACTTGGGTGACCAAG 1321
QY 1234 AAGCGGAGGAGGCCCTTCTTGGAGGAGAGCACTGCGAGCGTGTGGAGGTGACAGAG 1293
DB 1322 AAGCGGAGGAGGCCCTTCTTGGAGGAGAGCACTGCGAGCGTGTGGAGGTGACAGAG 1381
QY 1294 GGGAGGTTAAGAACTCAGTCAGGCTCATCCAGCTGGACCGGTCATCTCTGTTGAAG 1353
DB 1382 GAGAGTTTAAAGAACTCAGTCAGGCTCATCCAGCTGGACCGGTCATCTCTGTTGAAG 1441
QY 1354 TCCATGCTGAGGAGCGCTTCTTGGGAACCGCTTGGAGCCCGAGGACCGAGGAGGAAAG 1413
DB 1442 TCCATGCTGAGGAGCGCTTCTTGGGAACCGCTTGGAGCCCGAGGACCGAGGAGGAAAG 1501
QY 1414 CGATCCATGTCTGCTCCAGAAAACCTACTGCTGAAAGAAAGGTTTGGTGAAGGGGGCAAG 1473
DB 1502 CGATCCATGTCTGCTCCAGAAAACCTACTGCTGAAAGAAAGGTTTGGTGAAGGGGGCAAG 1561
QY 1474 AGCCCGAGCTCCCGCGCTCCAGGAGAGAGAGGCTGCATCC 1515
DB 1562 AGCCCGAGCTCCCGCGCTCCAGGAGAGAGAGGCTGCATCC 1603
```

```
RESULT 11
US-10-092-900A-53
; Sequence 53, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytex, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
```

```
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 53
; LENGTH: 1547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1529)
; US-10-092-900A-53

Query Match          96.1%; Score 1455.8; DB 13; Length 1547;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1492; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

QY      1  ATGGAGGGGGTCCAGCTCTGTGTCAGGATTCCTCGGGCAGAGCTGTGTAGAACGGGTG 60
Db      |||
QY      20  ATGGAGGGGGTCCAGCTCTGTGTCAGGATTCCTCGGGCAGAGCTGTGTAGAACGGGTG 79
Db      |||
QY      61  GCAGCATCATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCTACTAGAAAC 120
Db      |||
QY      80  GCAGCATCATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCTACTAGAAAC 139
Db      |||
QY      121  GGTGTGGACCCCCACACGGGCCAGAGCTGTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db      |||
QY      140  GGTGTGGACCCCCACACGGGCCAGAGCTGTGCTCTGTGATCCCTGGCAGTACTTCAAGA 199
Db      |||
QY      181  CTGCTCCAGCGCGGCTAGCCTCTCAGCAGAGAGCTTCCCTACAGAGCGGCCAGCA 240
Db      |||
QY      200  CTGCTCCAGCGCGGCTAGCCTCTCAGCAGAGAGCTTCCCTACAGAGCGGCCAGCA 259
Db      |||
QY      241  GGAAGCTATCTGGAGCGCAGGCTGGGCGCTTATGCGACGGGCGCTGCCAGCCACATCTCC 300
Db      |||
QY      260  GGAAGCTATCTGGAGCGCAGGCTGGGCGCTTATGCGACGGGCGCTGCCAGCCACATCTCC 319
Db      |||
QY      301  CCGCGGCGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db      |||

320  CCGCGGCGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 379
361  GACTCGTGCAGCTGAACAGGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 420
380  GACTCGTGCAGCTGAACAGGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 439
421  GTGGTGAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAGATTCCTTCC 480
440  GTGGTGAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAGATTCCTTCC 499
481  AAAAAGAGTGTACTGAAGCAGTATGGCTTTTCCACGTGCGCCCTCCCCCGAGAGGGTCCCAG 540
500  AAAAAGAGTGTACTGAAGCAGTATGGCTTTTCCACGTGCGCCCTCCCCCGAGAGGGTCCCAG 559
541  GTGCCCCAGGGAGGACCCAGCCAGAGCTGTGCGCCCTGGAGCGGGGTGTACAGAGATT 600
560  GCTGCCCGAGGGAGGACCCAGCCAGAGCTGTGCGCCCTGGAGCGGGGTGTACAGAGATT 619
601  GCATCTCTGAAGAAGCTGCACCAAGTGAATGTGTCAAACTGATCGAGGTCTCTGATGAC 660
620  GCATCTCTGAAGAAGCTGCACCAAGTGAATGTGTCAAACTGATCGAGGTCTCTGATGAC 679
661  CCAGCTGAGGACAACTCTTATTTGGTGTGTGACCTCTCGAGAAAGGGGGCCCTCATGGAA 720
680  CCAGCTGAGGACAACTCTTATTTGGTGTGTGACCTCTCGAGAAAGGGGGCCCTCATGGAA 736
721  GTGCGCTGTGACAAGCCCTTTCGAGAGGAGCAAGCTGCGCTCTACTTCGGGAGCTCATC 780
737  GTGCGCTGTGACAAGCCCTTTCGAGAGGAGCAAGCTGCGCTCTACTTCGGGAGCTCATC 796
781  CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTGCCAGGAGCATCAAGCCATCCAAC 840
797  CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTGCCAGGAGCATCAAGCCATCCAAC 856
841  CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGCACTTTGGGCTCAGCAACCAAGTTT 900
857  CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGCACTTTGGGCTCAGCAACCAAGTTT 916
901  GAGGGAGCAGCGCTCAGCTGTCCAGCAGCGGGGAGACCCAGCATTATGSCCCCCGAG 960
917  GAGGGAGCAGCGCTCAGCTGTCCAGCAGCGGGGAGACCCAGCATTATGSCCCCCGAG 976
961  GCATTTCTGATTCGGCCAGAGCTTCAGTGGAGAGGCTTGGATGTATGGGCCACTGGC 1020
977  GCATTTCTGATTCGGCCAGAGCTTCAGTGGAGAGGCTTGGATGTATGGGCCACTGGC 1033
1021  GTCACTGTGTACTGCTTTGTCTATGGGAAGTGCCTTCAATCGAAGATTTCAATCTGGCC 1080
1034  GTCACTGTGTACTGCTTTGTCTATGGGAAGTGCCTTCAATCGAAGATTTCAATCTGGCC 1093
1081  CTCCACAGGAAGATCAAGATCAGCCGCTGTGTTCTCTGAGAGCCAGAAATCAGCGAG 1140
1094  CTCCACAGGAAGATCAAGATCAGCCGCTGTGTTCTCTGAGAGCCAGAAATCAGCGAG 1153
1141  GAGCTCAAGGACCTCATCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTTGGGGTG 1200
1154  GAGCTCAAGGACCTCATCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTTGGGGTG 1213
1201  CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGGAGCCCTTCTCTCGGAG 1260
1214  CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGGAGCCCTTCTCTCGGAG 1273
1261  GAGGAGCATCTGAGCGTGTGGAGGTGACAGAGGGGGAGTTAAGAATCAGTCAGGCTC 1320
1274  GAGGAGCATCTGAGCGTGTGGAGGTGACAGAGGGGGAGTTAAGAATCAGTCAGGCTC 1333
1321  ATCCCCAGCTGACCAAGCTGTCTTGGTGAAGTCCATCTGAGGAAGGCTTCTTTGGG 1380
1334  ATCCCCAGCTGACCAAGCTGTCTTGGTGAAGTCCATCTGAGGAAGGCTTCTTTGGG 1393
1381  AACCCGCTTGAGCCCCCAGGACGGAGGAGCGGATCCATCTGCTCAGGAACCTTA 1440
```

Db 1394 AACCCGTTTGAGCCCAAGCACGAGGAGGAAAGAGGATCCATGTCTGTCCAGGAAACCTA 1453
Qy 1441 CTGTGTAAGAGAGGTTTGGTGAAGGGGCAAGAGCCACAGAGCTCCCGGGCTCCAGGAA 1500
Db 1454 CTGTGTAAGAGAGGTTTGGTGAAGGGGCAAGAGCCACAGAGCTCCCGGGCTCCAGGAA 1513
Qy 1501 GACGAGGCTGCATCC 1515
Db 1514 GACGAGGCTGCATCC 1528

RESULT 12

US-10-092-900A-57

; Sequence 57, Application US/10092900A

; Publication No. US20040043382A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Taupier Jr., Raymond J.

; APPLICANT: Pena, Carol E.A.

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Ji, Weizhen

; APPLICANT: Gorman, Linda

; APPLICANT: Miller, Charles E.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gangolli, Esha A.

; APPLICANT: Vernet, Corine A.M.

; APPLICANT: Guo, Xiaojia Sasha

; APPLICANT: Tchernev, Velizar T.

; APPLICANT: Fernandes, Elma R.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Liu, Yi

; APPLICANT: Anderson, David W.

; APPLICANT: Spaderna, Steven K.

; APPLICANT: Carterton, Elina

; APPLICANT: Leite, Mario W.

; APPLICANT: Zhong, Haihong

; APPLICANT: Alsbrook, John P.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieget, Daniel K.

; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-290C

; CURRENT APPLICATION NUMBER: US/10/092,900A

; CURRENT FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: USN 60/274,322

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USN 60/283,675

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: USN 60/338,092

; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: USN 60/274,281

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USN 60/274,191

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USN 60/325,681

; PRIOR FILING DATE: 2001-09-27

; PRIOR APPLICATION NUMBER: USN 60/304,354

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: USN 60/279,995

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: USN 60/294,899

; PRIOR FILING DATE: 2001-05-31

; PRIOR APPLICATION NUMBER: USN 60/287,424

; PRIOR FILING DATE: 2001-04-30

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 768

; SEQ ID NO 57
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(1681)
US-10-092-900A-57

Query Match

91.2%; Score 1381.4; DB 13; Length 1725;

Best Local Similarity 92.6%; Pred. No. 0;

Matches 1509; Conservative 0; Mismatches 6; Indels 114; Gaps 1;

Qy 1 ATGAGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 60
Db 52 ATGAGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 111
Qy 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCCTTACTAGAAAC 120
Db 112 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCCTTACTAGAAAC 171
Qy 121 GGTGTGACCCGCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 172 GGTGTGACCCGCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 231
Qy 181 CTGCTCCAGCCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db 232 CTGCTCCAGCCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 291
Qy 241 GGAAGCTATCTGAGGCGCAGGGCTGGGCCCTTATGCGACGGGGCTGCCAGCACATCTCC 300
Db 292 GGAAGCTATCTGAGGCGCAGGGCTGGGCCCTTATGCGACGGGGCTGCCAGCACATCTCC 351
Qy 301 CCCGGGCGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGGCCATCTCAGANTGCGAG 360
Db 352 CCCGGGCGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGGCCATCTCAGANTGCGAG 411
Qy 361 GACTGCGTGAGCTGAAACCAAGTCAAGCTGCAGAGTGAGATTGGCAAGGGTGCTACGGT 420
Db 412 GACTGCGTGAGCTGAAACCAAGTCAAGCTGCAGAGTGAGATTGGCAAGGGTGCTACGGT 471
Qy 421 GTGTGAGGCTGGCCTTCAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db 472 GTGTGAGGCTGGCCTTCAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 531
Qy 481 AAAAAGAAAGTTACTGAAGCAGTATGGCTTTCCAGCTGCCCTCCCGAGAGGGTCCAG 540
Db 532 AAAAAGAAAGTTACTGAAGCAGTATGGCTTTCCAGCTGCCCTCCCGAGAGGGTCCAG 591
Qy 541 GCTGCCCGAGGAGGACCCAGCAAGCAGCTGTGCTCCCTGGAGCGGGTGTACAGAGAGATT 600
Db 592 GCTGCCCGAGGAGGACCCAGCAAGCAGCTGTGCTCCCTGGAGCGGGTGTACAGAGAGATT 651
Qy 601 GCCATCTGAAAGAGCTGAGCCAGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 660
Db 652 GCCATCTGAAAGAGCTGAGCCAGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 711
Qy 661 CCAGCTGAGGACCAACCTCTATTG----- 684
Db 712 CCGGCTGAGGACCAACCTCTATTG----- 771
Qy 685 ----- 684
Db 772 TCAACAAATATGCCAAGTCCCACTCCCTGCTTCCCTCTGAGCAGCAAGACAGTGGATCC 831
Qy 685 -----GTGTTTGACCTCTGAGAAAGGGGCCCTCATGGAAGTGCC 726
Db 832 ACSTGGGCTCGGGCTCAGTGTGTTGACCTCTGAGAAAGGGGCCCTCATGGAAGTGCC 891
Qy 727 TGTGACAAGCCCTTCTCGGAGGAGCAGCTGCTCTACCTGCGGACGTCATCTCTGGC 786
Db 892 TGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCTCTACCTGCGGAGCGTCATCTCTGGC 951

```
QY 787 CTCGACTACTTGCACATGCGAGAGATCGTCCACAGGAGCATCAAGCCATCCACCTGCTC 846
Db 952 CTCGAGTACTTGCACATGCGAGAGATCGTCCACAGGAGCATCAAGCCATCCACCTGCTC 1011
QY 847 CTGGGGGATGATGGGCACGTGAAGATCGCGAGCTTTGGGCTCAGCAACACAGTTTGGGG 906
Db 1012 CTGGGGGATGATGGGCACGTGAAGATCGCGAGCTTTGGGCTCAGCAACACAGTTTGGGG 1071
QY 907 AAGGAGCTCAGTGTGACGACGGGGGAAACCCAGCATTCATGGCCCCCGAGGCCATT 966
Db 1072 AAGGAGCTCAGTGTGACGACGGGGGAAACCCAGCATTCATGGCCCCCGAGGCCATT 1131
QY 967 TCTGATTCGGGCGACAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGCGCTCAAG 1026
Db 1132 TCTGATTCGGGCGACAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGCGCTCAAG 1191
QY 1027 TTGTACTGCTTTGTCTATGGGAAGTGCCTCATTCATCGAGATTCATCTGCGCTCCAC 1086
Db 1192 TTGTACTGCTTTGTCTATGGGAAGTGCCTCATTCATCGAGATTCATCTGCGCTCCAC 1251
QY 1087 AGGAGATCAAGAATGAGCCCGTGTCTTCTGAGGAGCCAGAAATCAGCGAGGAGTTC 1146
Db 1252 AGGAGATCAAGAATGAGCCCGTGTCTTCTGAGGAGCCAGAAATCAGCGAGGAGTTC 1311
QY 1147 AAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTTGGGTGCCAGAC 1206
Db 1312 AAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTTGGGTGCCAGAC 1371
QY 1207 ATCAAGTTGACCCCTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTTCCGAGGAGGAG 1266
Db 1372 ATCAAGTTGACCCCTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTTCCGAGGAGGAG 1431
QY 1267 CACTCAGAGTGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1326
Db 1432 CACTCAGAGTGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1491
QY 1327 AGCTGACCAACGCTGATCCTGTGAGTCCATGCTGAGGAAGCGTTTCCTTTGGGAACCG 1386
Db 1492 AGCTGACCAACGCTGATCCTGTGAGTCCATGCTGAGGAGCGTTTCCTTTGGGAACCG 1551
QY 1387 TTTGAGCCCCAGGACGAGGAGGAAGCGATCGATCGTCTGCTCAGGAAACCTACTGTGTG 1446
Db 1552 TTTGAGCCCCAGGACGAGGAGGAAGCGATCGATCGTCTGCTCAGGAAACCTACTGTGTG 1611
QY 1447 AAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAAGACGAG 1506
Db 1612 AAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAAGACGAG 1671
QY 1507 GCTGCATCC 1515
Db 1672 GCTGCATCC 1680
```

RESULT 13

```
US-10-316-124-7
; Sequence 7, Application US/10316124
; Publication No. US20030152574A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO TREAT
; FILE REFERENCE: CARDIOVASCULAR DISEASE USING 1419, 58765, AND 2210
; CURRENT APPLICATION NUMBER: US/10/316,124
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/339,995
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Fastseq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4427
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-10-316-124-7
Query Match 43.0%; Score 651.2; DB 15; Length 4427;
Best Local Similarity 71.3%; Pred. No. 4.3e-181;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;
QY 267 GCTTATATGACAGGGGCTCCAGCCACATCTCCCGCGGCTGCGGGAGGCCACCAT 326
Db 360 GCTTATATGACAGGGGCTCCAGCCACATCTCCCGCGGCTGCGGGAGGCCACCAT 419
QY 327 CGAGTCCCAACCGTGGCCATCTCAGATCAGAGGAGTCTGCTGAGCTGAAACAGTACAA 386
Db 420 GGAGTCTCAACCGTGGCCATCTCAGATCAGAGGAGTCTGCTGAGCTGAAACAGTAC 479
QY 387 GCTGAGATGAGATTGGCAAGGGTGCCTACGGTGTGTTGAGGCTGGCCCTACAAAGAA 446
Db 480 CTTGAGGATGAAATTTGGAAAGGGCTCTTATGTTGCTGCTCAAGTTGGCTTACAA 539
QY 447 TGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTTACTGAAGCAGTAT 506
Db 540 TGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTTACTGAAGCAGTAT 599
QY 507 CTTTCCAGCTCGCCCTCCCGAGAGGGTCCAGAGCTGCCAGGAGGAGGAGGAGGAGG 566
Db 600 CTTTCCAGCTCGCCCTCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 659
QY 567 GCTGCTGCCCCCTGGAGCGGTGTACAGGAGATTTGCCATCTCTGAAGAGCTGACAC 626
Db 660 CAGGGGCCCCCATTTGAGCAGGTGTACAGGAAATTTGCCATCTCTGAAGAGCTGAC 719
QY 627 GAATGTGTGCAAACTGATCGAGGTCTGTGATGATCACCAGCTGAGGACCACTCTATT 686
Db 720 GAATGTGTGCAAACTGATCGAGGTCTGTGATGATCACCAGCTGAGGACCACTCTATT 779
QY 687 GTTGAACCTCTGAGAAAGGGGCGGTCTATGAAGTGCCTCTGACAGCCCTTCTCGGA 746
Db 780 GTTGAACCTCTGAGAAAGGGGCGGTCTATGAAGTGCCTCTGACAGCCCTTCTCGGA 839
QY 747 GAGCAAGCTCGCCCTTACTCTGGGAGCGTCTCTGGGCGCTCGAGTACTTGCACCTG 806
Db 840 AGACAGGCGCCCTTCTCTCTGAGGATCTGATCAAGGAGCTGAGTACTTGCACCTG 899
QY 807 GAAGATCTGTCCAAGGAGCATCAAGCCATCCAACTGCTCCTGGGGGATGATGGGAC 866
Db 900 GAAGATCTGTCCAAGGAGCATCAAGCCATCCAACTGCTCCTGGGAGAGATGGGAC 959
QY 867 GAAGATCTGTCCAAGGAGCATCAAGCCATCCAACTGCTCCTGGGGGATGATGGGAC 926
Db 960 GAAGATCTGTCCAAGGAGCATCAAGCCATCCAACTGCTCCTGGGGGATGATGGGAC 1019
QY 927 CAGCGGGGAACCCAGGATTCATGCCCCCGAGGCCATTTCTGATTCGGGCGAGCTT 986
Db 1020 CAGCGGGGAACCCAGGATTCATGCCCCCGAGGCCATTTCTGATTCGGGCGAGCTT 1079
QY 987 CAGTGGGAAGGCTTTGGATGTATGGGCCACTGCGGCTCAGCTTGTACTGTTGTCTAT 1046
Db 1080 CAGTGGGAAGGCTTTGGATGTATGGGCCACTGCGGCTCAGCTTGTACTGTTGTCTAT 1139
QY 1047 GAAATGCCCATTCATCGAGATTTCTCTGTGCCCCCTCCACAGGAGATCAAGAA 1106
Db 1140 CAAATGCCCATTCATCGAGATTTCTCTGTGCCCCCTCCACAGGAGATCAAGAA 1199
QY 1107 CAGTGGGATTTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGAGCTGATCCT 1166
Db 1200 CAGTGGGATTTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGAGCTGATCCT 1259
QY 1167 GTTAGCAAGAAATCCCGAGAGAGAAATTTGGGTGCCAGACATCAAGTTGACCCCT 1226
Db 1260 GCTGAGCAAGAAATCCCGAGAGAGAAATTTGGGTGCCAGACATCAAGTTGACCCCT 1319
QY 1227 GACCAAGAACGGGAGGAGGCCCTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGG 1286
Db 1320 CACGAGGATGGGGCGGAGGCCGTTGCCCCGCGAGGATGAGAACTGCAAGCTGGT 1379
```



```
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 4942
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-51

Query Match      43.0%; Score 651.2; DB 15; Length 4942;
Best Local Similarity 71.3%; Pred. No. 4.5e-181;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY 267 GCCTTATGCCACGGGCGCTGCCAGCCACATCTCCCCCGGGCGCTGGCGGAGGCGCCACCAT 326
DB 536 GCCCTACTCACCGTCAGCTCCCGCAGTCCCTCGCTCGCTGCCCGCGGGCGCAGAGT 595
QY 327 CGAGTCCCAACGAGTGGGCATCTCAGATGCAGAGGACTCGGTGACGTGAACAGTACAA 386
DB 596 GGAGTCTCACACGCTCTCCATCACGGGTATGCAGGACTGTGTGACGTGAATCAGTATAC 655
QY 387 GCTGCAGAGTGAATGGCAAGGTCCTACGCTGTGTGGTGGCTGGCTGGCTACAAAGAAAG 446
DB 656 CTGTAAGGATGAATTTGGAAGGGCTCCTATGGTGTGTCAGTGGCTTACATGAAAA 715
QY 447 TGAAGACAGACATATGCAATGAAAGTCCCTTTCCAAAAAGAAAGTTACTGAAAGCAGTATGG 506
DB 716 TGACAATACCTACTATGCAATGAAGTGTGTGTCCTCCAAAGAAAGTCAATCCGAGCGCG 775
QY 507 CTTTCCACGTGCGCTCCCGGAGAGGTCGCCAGGCTGCCAGGAGGAGCAGCCAAAGCA 566
DB 776 CTTTCCACGTGCGCTCCCGGAGGACCCCGGAGGCTCCCTGGAGGCTGCATCCAGCC 835
QY 567 GCTGCTGCCCTCGGAGCGGTGTACCAGGAGATTGCCATCTGAAAGAGCTGGACCAAGT 626
DB 836 CAGGGGCCCCATTGACGAGGTGTACAGGAATGTCATCTCCAAAGAGCTGGACCAAGT 895
QY 627 GAATGTGTCAAACCTGATCGAGTCTCGATGATGATGATGATGATGATGATGATGATGATGAT 686
DB 896 CAATGTGTGAAGCTGGTGGAGGTCTCGATGATGATGATGATGATGATGATGATGATGATGAT 955
QY 687 GTTGTGACTCTCTCAGAAAGGGGCGCTCATGGAAGTGGCTGTGACAAAGCCCTTCTCGGA 746
DB 956 GTTGAACATGGTCAACAAAGGGGCGGTGATGGAAGTGGCCACCCCTCAAAACCACTCTCTGA 1015
QY 747 GGAGCAAGCTCGCTCTACCTCGGGAGCGTCATCTCTGGGCGCTCGAGTCTTGACATGCCA 806
DB 1016 AGACAGGCGCGCTTCTACTTCTCAGGATCTGATCAAGAGCATCGAGTACTTACATACCA 1075
QY 807 GAAGATCGTCCACAGGAGCATCAAGCCATCCAACTGCTCTCGGGGGATGATGGGCAAGT 866
DB 1076 GAAGATCATCCACCGTGAATCAAACTCTTCCAACTCTTGGTGGGAGAGATGGGCAAT 1135
QY 867 GAAGATCGCGGCTTGGCGGTGAGCAACCAAGTGTGAGGGGAGCAGCGCTCAGCTGCCAG 926
DB 1136 CAAGATCGCTGACTTGGTGTGAGCAATGAATTCAGGGGAGTGCAGCGCTCTCTCCAA 1195
```

Search completed: July 25, 2004, 14:06:56
Job time : 994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 04:23:48 ; Search time 57 Seconds
(without alignments)
852.223 Million cell updates/sec

Title: US-10-690-617-2

Perfect score: 2634

Sequence: 1 MEGGPAVCCQDPRAELVERV.....FEGGKSPQLPGVQEDRAAS 505

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2452	93.1	505	2 A57156	Ca2+/calmodulin-de
2	1494	56.7	587	2 JCS5669	Ca2+/calmodulin-de
3	1472.5	55.9	417	2 JE0191	calcium/calmodulin
4	1043	39.6	357	2 T37317	probable Ca2+/calm
5	620	23.5	652	2 S38666	serine/threonine-p
6	586	21.5	1142	2 S50632	protein kinase PAK
7	512	19.4	382	2 T47464	serine/threonine-p
8	507.5	19.3	560	2 S57252	probable serine/th
9	468	17.8	1246	2 T51085	related to protein
10	448	17.0	440	2 T14736	probable serine/th
11	437	16.6	440	2 T14735	probable serine/th
12	436.5	16.6	511	1 A56009	serine/threonine-s
13	434	16.5	461	2 T14822	probable serine/th
14	429.5	16.3	512	2 T07788	probable serine/th
15	426.5	16.2	512	1 JC1446	serine/threonine-s
16	423.5	16.1	504	2 T10449	probable serine/th
17	418.5	15.9	472	2 B90100	SNP-related kinase
18	416.5	15.8	441	2 E85362	hypothetical prote
19	415.5	15.8	441	2 C84667	probable protein k
20	413.5	15.7	435	2 E84707	probable protein k
21	412.5	15.7	512	2 T52633	serine/threonine-s
22	410	15.6	389	2 S52242	protein kinase (EC
23	409.5	15.5	1192	2 T18611	probable serine/th
24	409	15.5	445	2 T50802	serine/threonine p
25	406.5	15.4	469	2 B84644	probable protein k
26	406	15.4	480	2 A86427	probable serine/th
27	405.5	15.4	504	2 T07415	probable serine/th
28	403.5	15.3	633	1 A26030	serine/threonine-s
29	403	15.3	576	2 T41587	probable carbon ca

ALIGNMENTS

RESULT 1

A57156

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - rat
N:Alternate names: Cam-kinase kinase; microtubule-associated protein 2 kinase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000
C:Accession: A57156; J04789
R:Tokumitsu, H.; Emslen, H.; Soderling, T.R.
J. Biol. Chem. 270, 19320-19324, 1995
A:Title: Characterization of a Ca(2+)/calmodulin-dependent protein kinase cascade. Molec
A:Reference number: A57156; MUID:95370263; PMID:7642608
A:Accession: A57156
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-505 <TK>
A:Cross-references: GB:I42810; NID:g986940; PIDN:AAC42070.1; PID:g986941
R:Okuno, S.; Kitani, T.; Fujisawa, H.
R. Biochem. 119, 1176-1181, 1996
A:Title: Evidence for the existence of Ca2+/calmodulin-dependent protein kinase IV kinase
A:Reference number: JC4789; MUID:96425004; PMID:8827455
A:Accession: JC4789

Query Match 93.1%; Score 2452; DB 2; Length 505;

Best Local Similarity 93.1%; Pred. No. 3 3e-101;

Matches 470; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1 MEGGPAVCCQDPRAELVERVAAIDVTHLEADGGPEPTNGVDPPPPRAAASVIGSTGR 60

DB 1 MERSPAVCCQDPRAELVERVAAISVAHLEAESEGPASNGVDPPSPRAAASVIGSASR 60

QY 61 LIPAPPSLSARKLSQERPAAGSYLEAQAGPYATGASHISPAWRPRTTIESHHVAISDAE 120

DB 61 PTFVPSLSARKFSQERPAESCLEAQGPYSTGSPASHMSPAWRPRTTIESHHVAISDTE 120

QY 121 DCVQLNQYKLSQSEIGKGVVRLAYNESDHYAMKVLKSKKLLKQYGFPRPPRGSQ 180

DB 121 DCVQLNQYKLSQSEIGKGVVRLAYNERDRHYAMKVLKSKKLLKQYGFPRPPRGSQ 180

QY 181 AAQGGPAKOLLPLERVYQEIALLKKLDHVNKKLIEVLDDPAEDNLVLFVLLRKGPNME 240

DB 181 APQGGPAKOLLPLERVYQEIALLKKLDHVNKKLIEVLDDPAEDNLVLFVLLRKGPNME 240

QY 241 VPCDPKPFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300

GIN4 protein - yea
serine/threonine-s
protein H39E23.1
probable serine/th
probable serine/th
FOG2 protein - yea
protein T25K16.13
myosin-light-chain
probable protein k
Ca2+/calmodulin-de
hypothetical prote
protein kinase (EC
hypothetical prote
p698g3 protein - A
SNF1-related prote
serine/threonine-s

Db 241 VPCDKPPEQARLYLRDIILGLEVLHCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
Qy 301 EGNDAQLSSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCFPIDDFTLA 360
Db 301 EGNDAQLSSTAGTAPFAPEAISDTGQSFSGKALDVWATGVTLYCFVYGKCFPIDEYILA 360
Qy 361 LHRKIKNEPVFPPEPEISEELKDLILKMLDKNPETRICVPDIKILHPWTKNGEPELPSE 420
Db 361 LHRKIKNEAVFPPEPEISEELKDLILKMLDKNPETRICVSDIKILHPWTKHGEPELPSE 420
Qy 421 BEHCSVVVEGEVKNVRLIPSWTIVLVKSMKRSGFNGPFPQARREERSMSAPGNL 480
Db 421 BEHCSVVVEEVEKNVSKLIPSWTIVLVKSMKRSGFNGPFPQARREERSMSAPGNL 480
Qy 481 LVKEGFGEGGKSPGLPGVQDEAAS 505
Db 481 LLKEGCGEGGKSPGLPGVQDEAAS 505

RESULT 2

JC5669
Ca2+/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change 24-Sep-1999
C:Accession: JC5669; PC4493
R:Kitani, T.; Okuno, S.; Fujisawa, H.
J. Biochem. 122, 243-250, 1997
A:Title: Molecular cloning of Ca2+/calmodulin-dependent protein kinase beta.
A:Reference number: JC5669; MUID:97420710; PMID:9276695
A:Accession: JC5669
A:Molecule type: mRNA
A:Residues: 1-587 <KIT>
A:Cross-references: GB:AB018081; NID:g3702720; PIDN:BAA33524.1; PID:d1034490; PID:g3702720
A:Experimental source: cerebellar
A:Accession: PC4493
A:Molecule type: protein
A:Residues: 425-501 <KI2>
A:Experimental source: cerebellar
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:162-445/Domain: protein kinase homology <KIN>
F:170-194/Domain: ATP-binding #status predicted <ATP>
F:480-493/Domain: calmodulin-binding #status predicted <CAB>

Query Match 56.7%; Score 1494; DB 2; Length 587;
Best Local Similarity 64.5%; Pred. No. 4.7e-59;
Matches 283; Conservative 75; Mismatches 55; Indels 26; Gaps 3;

Qy 64 ARPSLSARKLSQRRPAGSVLEAAGP-----YATGPASHISPRAWR 105
Db 89 SRSLSGGKNSLQR-----SOGGPASSSSLDNRCICPSLSYSPASSPQSPPR 141
Qy 106 RPTIESHHVAISDAEDCVQLNQYKIQSEIGKAGYGVVRLAYNESDRHYAMKVLKSKLL 165
Db 142 RPTVESHVSVITGLQDCVQLNQYTLKDEIGKSGYGVVKLAYNENDNTYAMKVLKSKLI 201
Qy 166 KQYGFPPRRPRGSOAQQGPAKQLPLERVYQIAILKLDHVNVLKLEVLDDPAEDN 225
Db 202 RQAGFPFRPPRRPRPAGGCIQPRGPIEQVYQIAILKLDHFNVLKLEVLDDPDNEH 261
Qy 226 LYLVDLRLKGPVMEVPCDKPFSBEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGD 285
Db 262 LYVVFELVNQGPVMEVTLKPLSEDAQRFYFDLKGLEYLHCQKIVHRDIKPSNLLVGE 321
Qy 286 DGHVKIADFGVSNQFEGNDAQLSSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYC 345
Db 322 DGHKIADFGVSNFEGKSDALLSNTVGTAPFAPEAISLSETRKIFSGKALDVWAMGVTLYC 381
Qy 346 FVYGKCFPIDDFTIALHRKIKNEPVFPPEPEISEELKDLILKMLDKNPETRICVPDIK 405
Db 382 FVFGCCPFMDRIMCLHSKIKSQALEFPDQDIAEDLKDLITRMLDKNPESRIVVPEIKL 441

Qy 406 HPWTKNGEPLPSEEHCSVVVEGEVKNVRLIPSWTIVLVKSMKRSGFNGPFP 465
Db 442 HPWTRREGAELPSEDEDNCTLVETEVEEVNSVKHIPSLATVILVKTMRKSRSGNPF 500
Qy 466 QARREERSMSAPGNLLVKE 484
Db 501 GSRREERSLSAPGNLLTKK 519

RESULT 3

JE0191
calcium/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) - human
X:Alternate names: CaMKK
C:Species: Homo sapiens (man)
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 16-Dec-1998
C:Accession: JE0191
R:Hsu, L.S.; Tsou, A.P.; Chi, C.W.; Lee, C.H.; Chen, J.Y.
Biomed. Sci. 5, 141-149, 1998
A:Title: Cloning, expression and chromosomal localization of human Ca2+/Calmodulin-depe
A:Reference number: JE0191
A:Accession: JE0191
A:Molecule type: mRNA
A:Residues: 1-417 <HSU>
C:Comment: This enzyme functions as a key intermediaries in Ca2+/calmodulin-driven sign
C:Genetics:
A:Gene: CaMKK
A:Map position: 12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase
F:47-330/Domain: protein kinase homology <KIN>

Query Match 55.9%; Score 1472.5; DB 2; Length 417;
Best Local Similarity 68.9%; Pred. No. 3e-58;
Matches 272; Conservative 72; Mismatches 50; Indels 1; Gaps 1;

Qy 90 PYATGPASHISPRARRPTTIESHVAISDAEDCVQLNQYKIQSEIGKAGYGVVRLAYNES 149
Db 11 PYSVPSSPQSPPRLPRPTVESHVSVITGMQDCVQLNQYTLKDEIGKSGYGVVKLAYNEN 70
Qy 150 EDRHYAMKVLKSKLLKQYGFPPRRPRGSOAQQGPAKQLPLERVYQIAILKLDHV 209
Db 71 DNTYAMKVLKSKLLKIQAGFPFRPPRRPRGTRPAPGGCIQPRGPIEQVYQIAILKLDHP 130
Qy 210 NVVKLIEVLDDPAEDNLVLDLRLKGPVMEVPCDKPFSBEQARLYLRDVLGLEYLHCQ 269
Db 131 NVVKLIEVLDDPDNEHLYVMFELVNQGPVMEVTLKPLSEDAQRFYFDLKGLEYLHYQ 190
Qy 270 KIVHRDIKPSNLLIGDDGHVKIADFGVSNQFEGNDAQLSSTAGTAPFAPEAISDSGQSF 329
Db 191 KIIHRDIKPSNLLVGEDGHIKIADFGVSNFEGKSDALLSNTVGTAPFAPEAISLSETRKIF 250
Qy 330 SGKALDVWATGVTLYCFVYGKCPPIDDFTIALHRKIKNEPVFPPEPEISEELKDLILKM 389
Db 251 SGKALDVWAMGVTLVCFVFGCCPFMDRIMCLHSKIKSQALEFPDQDIAEDLKDLITRM 310
Qy 390 LDKNPETRICVPDIKILHPWTKNGEPELPSEEHCSVVVEGEVKNVRLIPSWTIVIL 449
Db 311 LDKNPESRIVVPEIKLHPWTRHGAELPSEDEDNCTLVETEVEEVNSVKHIPSIAIVIL 370
Qy 450 VKSMKRKSRSGNPFEPQARREERSMSAPGNLLVKE 484
Db 371 VKTMRKSRSGNPF 404

RESULT 4

T37317
probable Ca2+/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) - Caenorhabditis
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37317
R:Edelman, A.M.; Mitchellhill, K.I.; Selbert, M.A.; Anderson, K.A.; Hook, S.S.; Stapleton
J. Biol. Chem. 271, 10806-10810, 1996
A:Title: Multiple Ca-calmodulin-dependent protein kinase kinases from rat brain.

A;Reference number: Z21682; MUID:96210015; PMID:8631893
A;Accession: T37317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-357 <EDE>
A;Cross-references: EMBL:AB016838; NID:g6624198; PIDN:BAA77824.3; PID:g6624199
A;Experimental source: strain Bristol N2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologous domain binding; phosphotransferase; protein kinase
F;Keywords: ATP; brain; calmodulin binding; phosphatase; protein kinase F;19-302/Domain: protein kinase homology <KIN>

Query Match 39.6%; Score 1043; DB 2; Length 357;
Best Local Similarity 57.2%; Pred. No. 1.9e+39;
Matches 198; Conservative 64; Mismatches 82; Indels 2; Gaps 2;

QY 119 AEDCVQLNOYNKLQSEIGKGAYGVVRLAYNESEDHRHYAMKVLSKKLLKOYGFFRRPPPRG 178
 :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 12 SESYLQNQRMLMEEELGGSGYGIWALYEENDKNYALKVLDRMKLLKNFACFRPPPRP 71
 ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 179 SOAQAGGPAPAKOLLPLERVQEIALKKLDHVNVKLVLELDDPAEDNLILVFLLRKGPV 238
 ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 72 NK-ENAAPSGLNRPLQLVOKEIALLKKLSHPNVKLVLEDNDPNLNLYMFVEFKESI 130
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 239 MEVPCKDPSEEQAIRYLVRVIILEGHLCOKIVRHDIKPNSLLLGDGHVKIAFGYSN 298
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 131 LEIPTDPKLUDEDTAWSYFRDLTCLGLELYHQKI VHRDI KP NS LLL SD I GK VI AD FG YSC 190
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 299 QPEGDAQSLSSTAGTPAFMAPEAISDSGSQP-SGKALDVATGTVTLYCFVVGKCPIFDIF 357
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 191 ELEGIDAFLSUTAGTGTAFAFAPEALTTEGANPHYSGRADQIWISGITLTLYAFVITGVTFVDNI 250
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 358 ILALHKRIKNEPVFPFEPEIPSELKDILLKMCDNPETRGVDPKLGHPWTNKGEPL 417
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 251 ITALHKIKNDPIVFPEAPILTSEAOLDILGMCLKDFCHRLMLEHVKHVTWRDGTVPM 310
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 418 PREEHCISVVETEGETEVONSURLIPSMTTLVTKSMLRKRFSGNPF 463
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 311 SSEQNCHELVTTVEEIEENCVRVIPRDITLIIVKMGHRCRFNGNF 356

RESULT 5
S58666
serine/threonine-protein kinase sspl - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 10-Dec-1999
C;Accession: S58666; T41270
R;Matsumoto, T.; Hirata, D.; Yanagida, M.; Toda, T.
EMBO J. 14, 3325-3338, 1995
A;Title: A novel protein kinase gene sspl(+) is required for alteration of growth polarity in fission yeast.
A;Reference number: S58666; MUID:95354651; PMID:7628434
A;Accession: S58666
A;Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A;Residues: 1-652 <NAT>
A;Cross-references: GI:D459882; NID:g1041184; PIDN:BAA08301.1; PID:d1008896; PID:g1041183
R;Purnelle, B.; Goffeau, A.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21983
A;Accession: T41270
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-652 <PUR>
A;Cross-references: EMU:AL049609; PIDN:CAB40783.1; GSPTDB:GN00068; SPDB:SPCC297.03
A;Experimental source: strain 972h-; cosmid c297
C;Genetics:
A;Gene: SPDB:SPCC297.03
A;Map position: 3
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologous domain binding motif
F;133-409/Domain: protein kinase homology <KIN>
F;141-149/Region: protein kinase ATP-binding motif

Query Match 23.5%; Score 620; DB 2; Length 652;

A;Map position: 5R
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: ATP; phosphotransferase; protein kinase
F;131-448/Domain: protein kinase homology <KIN>
F;139-147/Region: protein kinase ATP-binding motif

Query Match 21.5%; Score 566; DB 2; Length 1142;
Best Local Similarity 31.6%; Pred. No. 5.2e-18;
Matches 151; Conservative 82; Mismatches 135; Indels 110; Gaps 14;

QY 98 HISPRNWRPTIESHHVAISDAEDCVQ-----LNQYKLOSEIGKGGAYG 140
DB 86 HISSSLAKPTTTSSFCSSGSKNKVKETNRISLTYPVSKRVLTNTYIILKELGHGQG 145
QY 141 VRLAYNESEDRHYAMKVLK--KKLLKQYGFPRPPRPGSQAAQGGPAKQLLPLRVYQ 198
DB 146 KVKLRDILSKQLVAIKIVDRHEKKQKRFYFIK-----SSKISENDKIKR 191
QY 199 EIALKELKLDHNVVVKLIEVLDPAEDNLYLVFLLRKGPMVEVPCD-----KPFSE 249
DB 192 EIALMKCKHKHVVQVLEVLDDLSKRYLVLEVCSGEVVKWCPDQCMESDAKGPSLLSP 251
QY 250 EQARLYRDVILGLYHLCQKIVHRDIKPSNLLGDDGHVXIADFGVS-----NQPEGN 303
DB 252 QETREILRGVVLGLYHLCQKIVHRDIKPSNLLGDDGHVXIADFGVS-----NQPEGN 311
QY 304 DA---OLSSSTAGTPAFMAPE-AISDSG-----QSTSGKA-----LDVWATGVTLYC 345
DB 312 ESDLDELAKTVGTPAFAPEMWGEDAFTRYNUTKENLFRGSCISFMIDVWAGVTLYC 371
QY 346 FVYQKCFPIDFIALHARKIKNEVPVPEEPFIS-----EELKDLILKMLDK 392
DB 372 LLFCMLPFSSDFELKLFKIVNDLPKFTPEIKQSNKVKVSCSEYEMAKDILLKLEK 431
QY 393 NPETRIGVPDIKLPWVT-----KNGEPLPS-----EHEHGSVVEVTGEV 434
DB 432 NPQKMTIPAIKHPFVSMDPDHVPENDEKLLSVLEKLFQKFNQDQFEPISISKHEL 491
QY 435 KNSVRLIPSTVTILVKSM-----LRKRSFGNPFEPQARRERSMSAPCNLLAVKG 485
DB 492 KNAVSGVKKIKESVLKSIPLKOPSDLSNKNYLHPTTTRGRGD-----ANVIVSG 543

RESULT 7
T47464
serine/threonine-protein kinase-like protein - Arabidopsis thaliana
N;Alternate names: protein T14D3.180
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47464
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A;Reference number: 224467
A;Accession: T47464
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-382 <JOR>
A;Cross-references: EMBL:AL138649
A;Experimental source: cultivar Columbia; BAC clone T14D3
C;Genetics:
A;Map position: 3
A;Introns: 86/3; 111/3; 126/3; 151/3; 181/3; 222/3; 251/3; 312/3; 339/1
A;Note: T14D3.180

Query Match 19.4%; Score 512; DB 2; Length 382;
Best Local Similarity 37.0%; Pred. No. 4.4e-16;
Matches 119; Conservative 56; Mismatches 107; Indels 40; Gaps 9;

QY 107 PTIESHHVAISDAEDCVQ--LNQYKLOSEIGKGGAYGVRLAYNESEDRHYAMKVLKSKLL 165
DB 75 PVKETNKLTRGEDEGNKNTNEFVRERKIGSGYGVLYRSTVDDKHGAIKAFHKSLS 134
QY 166 KQYGFPRPPRPGSQAAQGGPAKQLLPLRVYQEIALLKLDHNVVVKLIEVLDPAEDN 225

A;Map position: 5R
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: ATP; phosphotransferase; protein kinase
F;131-448/Domain: protein kinase homology <KIN>
F;139-147/Region: protein kinase ATP-binding motif

Query Match 19.3%; Score 507.5; DB 2; Length 560;
Best Local Similarity 35.2%; Pred. No. 1e-15;
Matches 128; Conservative 58; Mismatches 115; Indels 63; Gaps 10;

QY 125 LNQYKLOSEIGKGGAYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRPGSQAAQ 184
DB 47 LNNFEILATLNGQYGVKVLARDLGTALVAIKILNR-----FEKR-----S 88
QY 185 GPAKQL-LPLRVYQEIALLKLDHNVVVKLIEVLDPAEDNLYLVFLLRKGPMVEVPCD 238
DB 89 GYSLQLKVENPRYNOQIEVNMKCHENNVVELYILNDPESTKVLVLEYSRGVFKWCP 148
QY 239 --MEVPCDKP--FSEQARLYRDVILGLYHLCQKIVHRDIKPSNLLGDDGHVXIAD 294
DB 149 NKMEIKAVGPSILTFQOSRKVLDVVGLEYLHSGIHTHRDIKPSNLLISNGTVKLSD 208
QY 295 GV-----SNOPEGNDAQL--SSTAGTPAFMAPEAISDSGSGSKALDVWATGVTL 345

RESULT 8
S57252
probable serine/threonine-specific protein kinase (EC 2.7.1.1-) YGL179c - yeast (Sacchar
N;Alternate names: protein BIE560; protein G1618
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
C;Accession: S57252; S64196; S56168
R;Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Bruschi, C.V.
Yeast 11, 767-774, 1995
A;Title: The DNA sequence of a 7941 bp fragment of the left arm of chromosome VII of Sa
ation and a putative serine/threonine protein kinase gene.
A;Reference number: S57252; MUID:95397594; PMID:7668046
A;Accession: S57252
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-560 <COG>
A;Cross-references: EMBL:X83690; NID:g794143; PIDN:CAA58659.1; PID:g794144
R;Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64183
A;Accession: S64196
A;Molecule type: DNA
A;Residues: 1-560 <BRU>
A;Cross-references: EMBL:Z72701; NID:g1322788; PIDN:CAA96891.1; PID:e243787; PID:g13227
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:TOS3
A;Cross-references: SGD:S0003147
A;Map position: 7L
A;Note: YGL179c
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; transmem
F;48-344/Domain: protein kinase homology <KIN>
F;56-64/Region: protein kinase ATP-binding motif
F;256-272/Domain: transmembrane #status predicted <TM>

Query Match 19.3%; Score 507.5; DB 2; Length 560;
Best Local Similarity 35.2%; Pred. No. 1e-15;
Matches 128; Conservative 58; Mismatches 115; Indels 63; Gaps 10;

QY 125 LNQYKLOSEIGKGGAYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRPGSQAAQ 184
DB 47 LNNFEILATLNGQYGVKVLARDLGTALVAIKILNR-----FEKR-----S 88
QY 185 GPAKQL-LPLRVYQEIALLKLDHNVVVKLIEVLDPAEDNLYLVFLLRKGPMVEVPCD 238
DB 89 GYSLQLKVENPRYNOQIEVNMKCHENNVVELYILNDPESTKVLVLEYSRGVFKWCP 148
QY 239 --MEVPCDKP--FSEQARLYRDVILGLYHLCQKIVHRDIKPSNLLGDDGHVXIAD 294
DB 149 NKMEIKAVGPSILTFQOSRKVLDVVGLEYLHSGIHTHRDIKPSNLLISNGTVKLSD 208
QY 295 GV-----SNOPEGNDAQL--SSTAGTPAFMAPEAISDSGSGSKALDVWATGVTL 345

Db 209 GVAMSTATGNTIQQSHHQLLKSRALGTPAFAPFELCSTEXKEYSCSSAIDWISLGVIIYC 268
Qy 346 FVYKCPDIDFILALHKKIKNEPVVPEEPEISE-----ELKDLILKMLDKN 393
Db 269 LFLGKLPMNAGLELFDISIINKPLFFPSYEMLNAGATGIMTEBYTDKALLKKLQKD 328
Qy 394 PETRIGVDPDIKLHPVWTNKGEPPLPS-----BEEH-----CSVVEVTEGEVKNVSR 439
Db 329 PDKRIKLADIKVHPFMCHYGKSDAASVLTNLETFLKVSPESSCKRVELVSLPWNSSFA 388
Qy 440 LIPS 443
Db 389 SLDS 392

RESULT 9
T51085
related to protein kinase PAK1 [imported] - Neurospora crassa
N:Alternate names: protein B2A19.190
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51085
R:Schulze, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51085
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 <SCH>
A:Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.190
A:Experimental source: BAC clone B2A19; strain OR74A
C:Genetics:
A:Gene: NCSP:B2A19.190
A:Map position: 6
A:Introns: 95/2

Query Match 17.8%; Score 468; DB 2; Length 1246;
Best Local Similarity 21.5%; Pred. No. 1.1e-13;
Matches 165; Conservative 91; Mismatches 168; Indels 344; Gaps 21;

Qy 37 PTRNGVDP-----PPRAAASVIPGSTSLPLA-----RPSLSARKLSLQRPAGSVLE 85
Db 19 PTQHHPRPLHPLPP---MASQIPHTANSTPVSSPGLFSPSPRNSMSLQSPSE--- 72
Qy 86 AQAGPYATGPASHIPRAWRRPTIESHHVAIDAEQCVQ-----LNQYKLOSEIGKG 137
Db 73 -NTTAPTILHSPYLP-----LQSHKVRTHKANVEHDYTTGRKAINQYHIIELGEG 124
Qy 138 AYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRPGSQAAQGGPAKQLPLERVY 197
Db 125 MHGKVKLARNTQTGDNVAIKII-----PRFSKRRLGKVTAMSTQDKSK 168
Qy 198 QEIALTKLDHNVNKLTEVLDDPAEDNLVLYFD----- 231
Db 169 REIALTKIRHPNVVALLIEDDPELKYTYMVLHVELGVVVRKKGPHICAYERRQEQ 228
Qy 232 ----- 231
Db 229 REQLGALPDAREEVYLRFMKQAEKAKAHVARQAHNNNAEFWSLEFGAADDLDDPH 288
Qy 232 ----- 231
Db 289 ARSLGRDSSIPSFAGDWLGLSRVTSRAPRSTQSMKSISSRNTPOPSEPDPHASIASVPE 348
Qy 232 -----LLRKGPMVE----- 240
Db 349 VDEDEMETPHGSHINQDSAIEDSLFAPRENEHSLFKRSPMSADSTLSHMSSVDYNRVHD 408
Qy 241 -----VPCDKPFSEEQARLYLRDVLIGLEYLHCQKIVHRDIKPSNLLDGDGHVKA 292
Db 409 PFVDDOYSYVFC-----FTIDKARSAFRDVLIGLEYLHVEGVVHRDIKPNALLCTKDRVKIS 465

Qy 293 DFGVS-----NOFEGNDAQISSLTAGTAPMAPE-ALSDSQSQSPG 331
Db 466 DFGVSYFGRLRDGEDPEVSESEARDFD-NDLAKTAVGTAPFAPELCYTDTYDDRP 524
Qy 332 -----KALDQWATGVTLVCFVYCKCPDIDFIFALHHRKIKNEPVVFP----- 373
Db 525 QQPKITQIDVMSLUGVTLYCLIFARIPLAEDEWRMFKFIATEDIYIPRQRLRPVDPSTK 584
Qy 374 -----EEP-----EISEELKDLILKMLDKNPKETRIGVDPDIKLHPWTKN 412
Db 585 PDEKSLYTRVNRDPYRNDDDEPLYEEDNDLYDLLSKMLTKNPEKRIRLRDVKRHPWLKD 644
Qy 413 GEEPL-----PSEEHCSVVVTEGEVKNVRLIP-----SWTIVILVKSMLKRSFGN 461
Db 645 IDNVIAWLDDTDPSSRTAGRIQVDRDITQAV--VPLTFVERAKLAFKKTIVTKFTHRGD 702
Qy 462 PFEQARREERSMSAPGNLLVKEGFGEGGKPEL-----PGVQDEEAAS 505
Db 703 RSESVSSKRKATSSAAS-----SAESPAGVGIPTFGVDRGERKS 741

RESULT 10
T14736
probable serine/threonine kinase (EC 2.7.1.1-) SNFL2 - sorghum
N:Alternate names: SNF1 protein kinase homolog SNFL2
C:Species: Sorghum bicolor (sorghum)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14736
R:Amnen, F.; Stockhaus, J.
Plant Mol. Biol. 36, 529-539, 1998
A:Title: Characterization of a Sorghum bicolor gene family encoding putative protein kin.
A:Reference number: Z18177; MUID:98145442; PMID:9484448
A:Accession: T14736
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-440 <ANN>
A:Cross-references: EMBL:Y12465; NID:g2632253; PIDN:CAA73068.1; PID:g2632254
A:Experimental source: cultivar TX 430
C:Genetics:
A:Gene: SNFL2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match 17.0%; Score 448; DB 2; Length 440;
Best Local Similarity 29.5%; Pred. No. 3.3e-13;
Matches 119; Conservative 75; Mismatches 139; Indels 70; Gaps 11;

Qy 124 QLNQYKLOSEIGKGAYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRPGSQAAQ 183
Db 9 RVGKVELGRTIGEGTFAKVRFAKNTGETGEVPAIKLDRKVKLRH----- 52
Qy 184 GCPAKQLPLERVYQEIALLKKLDHNVNKLTEVLDDPAEDNLVLYFDLLRKGPMVPEVC 243
Db 53 -----KMVEQIKREISTMKLIIKHPNVVRIVEVMG--SKTKIYIVLEVVTGGLFDITIA 103
Qy 244 DK-PPSEQARLYLRDVLIGLEYLHCQKIVHRDIKPSNLLDGDGHVKADEGVS--NQF 300
Db 104 NHGRNREDEARRYFOQLINAVDYCHSRGVYHRDLAKPENLLDSYGNLKVDSDFGSLSSQ 163
Qy 301 EGNDQAQLSTAGTAPFAPMAPEAISDSQSFSGKALDQVWATGVTLVCFVYKCPDIDDFILA 360
Db 164 IKDGLLHTTCGTPNYVAPEVLEQDQ--YDGAMADLWSGVILFVLLAGYLPFEDSNLMT 221
Qy 361 LHRKIKNEPVVPEPEI SEELKDLILKMLDKNPKETRIGVDPDIKLHPWTKNKEEP---- 416
Db 222 LYKKISNAEYFP--PWTSPFAKRLLTRFLDPNPWTRITIPLEIDEDWFKKGKPEFDE 279
Qy 417 -----LPSEEHCSVVVTEGEVKNVRLIPSWTIVILVKSMLKRSFGNPF 463
Db 280 KYDTPLDDVDVAFNDSEEHVTEKKEEPVVLNAFELI-----SRSAGLNLGNLF 329
Qy 464 --EPQARREERSMS--APGNLLVKE-----GFCGEGGKSP 495

Db 330 DSEQEKKRETRFTSKCPKPKETVRKIEBAAKPLGFGVKKNYL 372

RESULT 11

Tl4735
A;Title: probable serine/threonine kinase (EC 2.7.1.1-) SNFL1 - sorghum
N;Alternate names: SNF1 protein kinase homolog SNFL1
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: Tl4735
R;Annen, F.; Stockhaus, J.
Plant Mol. Biol. 36, 529-539, 1998
A;Title: Characterization of a Sorghum bicolor gene family encoding putative protein kinases
A;Reference number: Z18177; MUID:9814542; PMID:9484448
A;Accession: Tl4735
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-440 <ANN>
A;Cross-references: EMBL:Y12464; NID:g2632251; PIDN:CAA73067.1; PID:g2632252
A;Experimental source: cultivar TX 430
C;Genetics:
A;Gene: SNFL1
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match 16.6%; Score 437; DB 2; Length 440;
Best Local Similarity 29.7%; Pred. No. 1e-12;
Matches 114; Conservative 72; Mismatches 136; Indels 62; Gaps 10;

QY 124 QLNQYKLOSEIGKAGYVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQAQ 183

Db 9 RVGKYLGRITIGETFAKVFARDVTGAEVAIKLDKVLKH----- 52

QY 184 GGPAKQLPLERYVQETAILKLDHVNVLKLEVLDDPAEDNLVLPDLRKGPVME-VP 242

Db 53 -----KWEQKREISTMKLTKHPNVRIEYVMG--SKTKIYVLEFATGELFORIV 103

QY 243 CDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDKPSNLLLGDDGHVKIADFGVS--NQF 300

Db 104 NHGRMREDEARRYFQLINADVICHSGVYHRDLKPENLLDLSVGNLKVSDFGLSALSQQ 163

QY 301 EGNDAQLSSTAGTAPFAPMAIPAISSQSGSFGKALDVWATGVTLVCFYVYKCPDIDDFILA 360

Db 164 MKDDGLLHTTCGTFTYVAPEVLEDOG--YDGAMADLWSCGVILFVILAGYLPEDSNLMT 221

QY 361 LHRKIKNEPVVPEEPEISEELKDLILKMLDKNPETRIGVDPDKLHPWTKNGEPP--- 416

Db 222 LYKKISNAETFP--PWTSPAKRLITRIIDPNMTIRITPEILEDEWFKKGYKRPEFDE 279

QY 417 -----LPSEEEHCSVVEVTEGEVKNVRLIPSWTIVILVKSMLRKRSFGNPF 463

Db 280 KYDTTLDDVDVAFNDSEHHVTEKKEEPVALNAFELI-----SNSAGNLGNLF 329

QY 464 --EPQARREERSMS--APGNLLVKE 484

Db 330 DSEQEKKRETRFTSKCPKPKETVRK 353

RESULT 12

A56009
serine/threonine-specific protein kinase (EC 2.7.1.1-) NPK5 - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C;Accession: A56009
R;Muranaka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A;Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae
A;Reference number: A56009; MUID:94217693; PMID:8164654
A;Accession: A56009
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-511 <MUR>

A;Cross-references: GB:D26602; NID:g496384; PIDN:BAA05649.1; PID:g496385
C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C;Superfamily: AMP-activated protein kinase; protein kinase homolog
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F;17-271/Domain: protein kinase homology <KIN>
F;25-33/Region: protein kinase ATP-binding motif
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 16.6%; Score 436.5; DB 1; Length 511;
Best Local Similarity 33.7%; Pred. No. 1.2e-12;
Matches 96; Conservative 64; Mismatches 94; Indels 31; Gaps 6;

QY 125 LNQYKLOSEIGKAGYVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQAQ 184

Db 16 LRNYKLGKTLGIGSGFKVIAEHTLTGKVAVKILNRKI----- 55

QY 185 GGPAKQLPLERYVQETAILKLDHVNVLKLEVLDDPAEDNLVLPDLRKGPVMEVPCD 244

Db 56 ---KNWEMEEKVREIKLRLFMHPIRLIYEVVETPSD--IYVMEYVKSGLDFDIVE 110

QY 245 K-PFSEEQARLYLRDVLGLEYLHCQKIVHRDKPSNLLLGDDGHVKIADFGVSNQFEGN 303

Db 111 KGRLOEARKPFQIISGVEYCHRNVMVHRDLKPENLLDLSKNVVKIADFGLSNIMRDG 170

QY 304 DAQLSSTAGTAPFAPMAIPAISSQSGSFGKALDVWATGVTLVCFYVYKCPDIDDFILALHR 363

Db 171 HF-LKTSQSGPNAAPEVI--SGKLYAGPEVDVMSCGVILYALLCGTLFPDDENIPNLFK 227

QY 364 KIKNEPVVPEEPEISEELKDLILKMLDKNPETRIGVDPDKLHPW 408

Db 228 KIKGMSILPS--HLSAGARDLIPRLIYVDPKRMRTIPEIRMPW 270

RESULT 13

Tl4822
probable serine/threonine protein kinase (EC 2.7.1.1-) SNFL3 - sorghum
N;Alternate names: SNF1 protein kinase homolog SNFL3
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C;Accession: Tl4822
R;Annen, F.; Stockhaus, J.
Bot. Acta 111, 137-142, 1998
A;Title: SNFL3 a protein kinase homolog of Sorghum bicolor with a high similarity to the
A;Reference number: Z18217
A;Accession: Tl4822
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-461 <ANN>
A;Cross-references: EMBL:Y14274; NID:el301468; PID:el301469
A;Experimental source: cultivar TX 430
C;Genetics:
A;Note: SNFL3
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase; protein kinase
F;10-266/Domain: protein kinase homology <KIN>

Query Match 16.5%; Score 434; DB 2; Length 461;
Best Local Similarity 29.4%; Pred. No. 1.4e-12;
Matches 111; Conservative 73; Mismatches 135; Indels 58; Gaps 11;

QY 118 DAEDCVOLNOYKLOSBIGKAGYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPR 177

Db 2 DERRTILMDRYEIGQLGQGNFAKVYFARNLTDGQSVAKMIDKDKITR-VGL----- 53

QY 178 GSQAAGGPAKQLPLERYVQETAILKLDHVNVLKLEVLDDPAEDNLVLPDLRKGP 237

Db 54 -----IVQIKREISIMRLVKHPNLQLFEVM--ASKSKIYFVLEYAKGGE 96

QY 238 VMEVPCDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDKPSNLLLGDDGHVKIADFGVS 297

Db 97 LFNKISKGKFSBVARRYFHQLISAVDYCHSRGVVHRDLKPENLLDENENLKVSDFGLS 156

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 01:20:58 ; Search time 53 Seconds
(without alignments)
496.140 Million cell updates/sec

Title: US-10-690-617-2
Perfect score: 2634
Sequence: 1 MEGGPAVCCQDRAELVERV.....FGEKGSPFLPGVQDEAAS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620	23.5	652	1 SSP1_SCHPO	P50526 schizosacch
2	566	21.5	1142	1 PAK1_YEAST	P38990 saccharomyc
3	507.5	19.3	560	1 KGS9_YEAST	P43637 saccharomyc
4	437	16.6	620	1 SNF1_CANAL	P52497 candida alb
5	432	16.4	619	1 SNF1_CANTR	O94168 candida tro
6	426.5	16.2	535	1 K110_ARATH	Q38997 arabidopsis
7	412.5	15.7	756	1 DCK1_MOUSE	Q91lm8 mus musculu
8	410	15.6	408	1 ST6L_XENLA	Q91819 xenopus lae
9	407	15.5	611	1 SNF1_CANGA	Q00372 candida gla
10	403.5	15.3	633	1 SNF1_YEAST	P06782 saccharomyc
11	403	15.3	576	1 SNF1_SCHPO	O74536 schizosacch
12	403	15.3	1142	1 GIN4_YEAST	Q12263 saccharomyc
13	402.5	15.3	433	1 DCK1_RAT	O08875 rattus norv
14	401.5	15.2	740	1 DCK1_HUMAN	O15075 homo sapien
15	397	15.1	432	1 STKB_XENLA	Q91604 xenopus lae
16	396	15.0	295	1 KMLC_DICDI	P25323 dictyosteli
17	394	15.0	648	1 KAPC_DICDI	P34099 dictyosteli
18	393	14.9	472	1 KCCG_HUMAN	Q13555 homo sapien
19	390.5	14.8	407	1 STK6_XENLA	Q91820 xenopus lae
20	389.5	14.8	552	1 AA2K_RAT	Q09137 rattus norv
21	389	14.8	527	1 KCCG_RAT	P11730 rattus norv
22	387.5	14.7	552	1 AA2K_HUMAN	P54646 homo sapien
23	387	14.7	661	1 ARK5_HUMAN	O60285 homo sapien
24	387	14.7	794	1 K111_HUMAN	Q8tdc3 homo sapien
25	386	14.7	476	1 LOK_DROME	O61267 drosophila
26	385.5	14.6	529	1 KCCG_MOUSE	Q923c9 mus musculu
27	384.5	14.6	533	1 KCCD_RAT	P15791 rattus norv
28	384.5	14.6	1037	1 KCC4_YEAST	P25389 saccharomyc
29	384	14.6	478	1 KCCA_HUMAN	Q9ucm7 homo sapien
30	384	14.6	543	1 CHK2_HUMAN	O96017 homo sapien
31	383.5	14.6	478	1 KCCA_RAT	P11275 rattus norv
32	383	14.5	433	1 STKB_HUMAN	Q15831 homo sapien
33	383	14.5	774	1 MRK2_MOUSE	Q05512 mus musculu

34	382.5	14.5	478	1 KCCA_MOUSE	P11798 mus musculu
35	382.5	14.5	499	1 KCCD_HUMAN	Q13557 homo sapien
36	382	14.5	460	1 CDS1_SCHPO	Q09170 schizosacch
37	378.5	14.4	664	1 KCCB_HUMAN	Q13554 homo sapien
38	378	14.4	404	1 KAPC_CABEL	P21137 caenorhabdi
39	378	14.4	542	1 KCCB_MOUSE	P28652 mus musculu
40	377	14.3	350	1 KAPA_MOUSE	P05132 mus musculu
41	377	14.3	481	1 AKT2_HUMAN	P31751 homo sapien
42	375.5	14.3	350	1 KAPA_HUMAN	P17612 homo sapien
43	375.5	14.3	550	1 AAK1_HUMAN	Q13131 homo sapien
44	375.5	14.3	623	1 NRC2_NEUCR	O42626 neurospora
45	375.5	14.3	891	1 KIN1_SCHPO	P22987 schizosacch

ALIGNMENTS

RESULT 1
SSP1_SCHPO
ID SSP1_SCHPO STANDARD; PRT; 652 AA.
AC P50526;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase sspl (EC 2.7.1.-).
GN SSPI OR SPCC297.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=95354651; PubMed=7628434;
RA Matsuoka T., Hirata D., Yanagida M., Toda T.;
RT "A novel protein kinase gene sspl is required for alteration of
RT growth polarity and actin localization in fission yeast.";
RL EMBO J. 14:3325-3338(1995).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.-J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
CC -!- FUNCTION: Involved in actin localization and thus in polarized
CC cell growth.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D45882; BAA0301.1; -.
CC EMBL; AL049609; CAB40783.1; -.
CC PIR; S58666; S58666.
CC HSP; Q63450; IA06.
CC GeneDB SPombe; SPCC297.03; -.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser Thr_pkin_AS.
CC InterPro; IPR002290; Ser Thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC DOMAIN 135 409 PROTEIN KINASE.
FT NP BIND 141 149 ATP (BY SIMILARITY).
FT BINDING 164 164 ATP (BY SIMILARITY).
FT ACT SITE 267 267 BY SIMILARITY.
SQ SEQUENCE 652 AA; 73992 MW; B1931E5EB75A85DA CRC64;

Query Match
Best Local Similarity 33.5%; Score 620; DB 1; Length 652;
Matches 168; Conservative 67; Mismatches 176; Indels 78; Gaps 14;

QY 52 SVTPGSTRLLPAPSLARKLSLQER--PAGSYLEAQAGPYATG-----PASH- 98
Db 30 SEVPSPVFPYVHKIAQIKTSDFRKNVSAGDYVIAPLGREGESLTHSWTFQPKGN 89

QY 99 -----ISPRAWRR-----PTTSHHVAISDAEDCVQLNOYKLOEIGKAY 139
Db 90 QRLYSNDFQEAQWKRLOEWGEVKTTKIRKDFRFSGRK---YINHYLEIKELGRGMH 146

QY 140 GVVRLAYNSEDRHYAMKVLKSKKLLKQYGFPRRPPRPGSQAAQGGPAKQLPLERVYQE 199
Db 147 GKVKLGRDTWTRELLAIKIIPKE-----RRP-----KLGRANASSQKEKVRRE 190

QY 200 IALLKLDHVVNVLKLEVLDDPAEDNLVLFVDFLLRKGPVWEVPCDPF-PSEQARLYLRD 258
Db 191 IALLKVCVHPNVVRUREVIDDPSSTKVLVLEYMSGGVEPWTDCDSFVLISSEARQYFRD 250

QY 259 VILGLEYLHCQKIVHRDKIPSNLLDGDGHVKIADFGVS-----NQFEGNDAQLSSTAGT 313
Db 251 VVLGLEYLHVQGIHRDIKPANLLNSNCVKLSDFGVSYIANAGLNENDVELAKTVGT 310

QY 314 PAFMAPEAL---SDSQSFGSKALDWATGVLYCFVYKCPIDDFIILALHRKIKNEPV 370
Db 311 PAFPAPELCTWLDLRPRPKISEAIDWALGVTLFCLLFGRCFENASMEYELFDKI VNERL 370

QY 371 VPPEPELSEELKDLILKMLDNPETRIGVDFDKLHPWTKNGEEP-----LPSEEHHCVS 426
Db 371 NTPSPDIDGEGRDLLKLLCKDPEORITLVEVKLHPWTLGDLGKWLQNTDPSTVSR 430

QY 427 VEVTEGEVKNVRLIPSWTTTIVLKSMLKRRFGNPF---EPQARPEERSMAGPGLLVK 483
Db 431 VEVSTDEVASALSLV-----GLRL-RKLGKLFRRFRPKARVFDSSSSVPDSSSIC 479

QY 484 EGFEGGGKS 492
Db 480 RPSSGNS 488
```

RESULT 2
PAK1_YEAST

```
ID PAK1_YEAST STANDARD; PRT; 1142 AA.
AC P38990;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase PAK1 (EC 2.7.1.-).
GN PAK1 OR YER129W OR SYGP-ORF45.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313264; PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mossdale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomycetes cerevisiae chromosome V.";
RL Nature 387:78-81(1997).
RN [2]
RP SEQUENCE OF 1-657 FROM N.A.
RC STRAIN=A364A;
RX MEDLINE=98000885; PubMed=9341678;
RA Hovland P.G., Tecklenberg M., Sclafani R.A.;
RT "Overexpression of the protein kinase Pak1 suppresses yeast DNA
polymerase mutations.";
RL Mol. Gen. Genet. 256:45-53(1997).
CC -!- FUNCTION: May function by modifying and partially stabilizing
CC thermolabile DNA polymerases, perhaps during DNA repair.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18916; AAC03227.1; -.
CC EMBL; U13398; AAC49840.1; ALT_TERM.
CC PIR; S50632; S50632.
CC GerMOnline; 139208; -.
CC SGD; S0000931; PAK1.
CC GO; GO:004672; F:protein kinase activity; IDA.
CC GO; GO:000261; P:DNA dependent DNA replication; IGI.
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR002290; Ser Thr_pkin_AS.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation.
FT DOMAIN 133 448 PROTEIN KINASE.
FT NP BIND 139 147 ATP (BY SIMILARITY).
FT BINDING 162 162 ATP (BY SIMILARITY).
FT ACT SITE 277 277 BY SIMILARITY.
FT CONFLICT 171 171 Q -> H (IN REF. 2).
FT CONFLICT 266 268 EYL -> DS (IN REF. 2).
SQ SEQUENCE 1142 AA; 126871 MW; 425D71B8340B3F8F CRC64;

Query Match
21.5%; Score 566; DB 1; Length 1142;
```

```
Best Local Similarity 31.6%; Pred. No. 3e-29;
Matches 151; Conservative 82; Mismatches 135; Indels 110; Gaps 14;

Qy 98 HISPRWRPTTIESHHVAISDAEDVQ-----LNOYKLOSEIGKGAYG 140
Db 86 HISSSLAKPTTTSFSGSGSKNKKVETNRISITYDPVSRKVLNYYEIIKELGHGQHG 145
Qy 141 VVRLAYNESEDRHYAMKVLK--KKLLKQYGFPRPPRPGSQAAQGGPAKOLLPLERYVQ 198
Db 146 KVKLARDILSKQLVAIKIVDRHEKKQKFFFIK-----SSKISENDKIKR 191
Qy 199 EIALKCLKDHNVVKLIVLDDPAEDNLYIFVDDLKPKPMEVPCD-----KPFSE 249
Db 192 EIALMKCHHHVVQVLEVDLRSKTKLYLVLEYSRGEVKWCPDCMESDAKGFSLLSF 251
Qy 250 EEARLYLRDVLGLLEYLHCQIKVHRDIKPSNLLGDDGHVKIADPGVS-----NQFEGN 303
Db 252 QETREILGVVGLLEYLHYQGIHHRDIKPAVLLISGDTGVKISDFGVSLAASSTNSDSS 311
Qy 304 DA----QLSSTAGTAPFAWPE--AISDSG-----QSFGKA-----LDVWATGVTLYC 345
Db 312 ESLDELELAKTVGTPAFAPWEMCLGEDAFTRYNLTKENLFRGSCISFMIDIWAVGVTLYC 371
Qy 346 FVYGKCPIDDFILALHKKIKNEPVVFPPEEIS-----PELKOLILKMLDK 392
Db 372 LLFGMLPFPDFELKLFKEKIVNDPLKFTFKETIQSNKVSQVSCBEEYEMAKDLLKLEK 431
Qy 393 NPETRIGVDPDKLHPWT-----KNGEELPS-----EEHCSEVVEVTEGEV 434
Db 432 NPQKMTIPALKQPFVSWDFHVDENDEKLLSSVLEQKLPQCQNTQFPFISIKHEL 491
Qy 435 KNSVRLIPSWTITVLVSKM-----LRKSGFNGFPFQAPAREERSMSAPGNLLYKEG 485
Db 492 KNAVSGVKIKESVLKSIPLKPSDLNKNKYLHPTETTRGRD-----ANVIVSEG 543

RESULT 3
KGS9 YEAST
ID KGS9 YEAST STANDARD; PRT; 560 AA.
AC P43637;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase YGL179C (EC 2.7.1.1.-).
GN YGL179C OR G1618.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=95397594; PubMed=7668046;
RA Coglievina M., Bertani I., Klina R., Zaccaria P., Bruschi C.V.;
RT "The DNA sequence of a 7941 bp fragment of the left arm of chromosome
RT VII of Saccharomyces cerevisiae contains four open reading frames
RT including the multicopy suppressor gene of the pop2 mutation and a
RT putative serine/threonine protein kinase gene.";
RL Yeast 11:767-774 (1995).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X83690; CAA58659.1; -
CC EMBL; Z72701; CAA96891.1; -
CC FIR; S57252; S57252.
CC HSSP; Q00534; 1B18.
```

```
GermOnline; 141227; -.
DR SGD; S0003147; TOS3.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase, 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
FT ATP-binding.
KW DOMAIN 50 344 PROTEIN KINASE.
FT NP BIND 56 64 ATP (BY SIMILARITY).
FT BINDING 79 79 ATP (BY SIMILARITY).
FT ACT SITE 189 189 BY SIMILARITY.
FT ACT SITE 189 189 BY SIMILARITY.
SQ SEQUENCE 560 AA; 62090 MW; EPPA0C5ED58B5466 CRC64;

Query Match 19.3%; Score 507.5; DB 1; Length 560;
Best Local Similarity 35.2%; Pred. No. 7.8e-26;
Matches 128; Conservative 58; Mismatches 115; Indels 63; Gaps 10;

Qy 125 LNOYKLOSEIGKGAYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRPGSQAAQG 184
Db 47 LNFEEILATLNGQYQKVKLARDLGTGALVAIKILNR-----FEKR-----S 88
Qy 185 GPAKQL-LPLERVQEIATLKKLDHNVVVKLIEVLDDPAEDNLYLVFOLLKGPV----- 238
Db 89 GYSIQLKVENPRVQIEVEMKRCHEENVVELYEILNDPESTKVYLVLVEYCSRGPKWCE 148
Qy 239 --MEVPCDKP--FSEEQARLYLRDVLGLLEYLHCQIKVHRDIKPSNLLGDDGHVKIADP 294
Db 149 NKMEIKAVGPSILATFQOSRKVVLVDVWSGLEYLHSGQITHRDIKPSNLLISSNGTVKISDF 208
Qy 295 GV-----SNQFSGDAQL--SSTAGTAPFAWPEAISDSQSFGSKALDVWATGVTLYC 345
Db 209 GVAMSTATGNTSIQSSHEQLKRALGTPAFAPFALCSTEKYECSSAIDVLSLGVTLYC 268
Qy 346 FVYGKCPIDDFILALHKKIKNEPVVFPPEEPISE-----ELKOLILKMLDKN 393
Db 269 LLFGKLPFNANSGLFLPDSILNKPLEFYSYEMLNGATSGITMEYTDKALLKLLQKD 328
Qy 394 PETRIGVDPDKLHPWTVKNGEELPS-----EEEH-----CSVVEVTEGEVKNVSR 439
Db 329 PDKRIKLADIKVHPFPMCHYGKSDAASVLTNLETFHELKVSPPSSCKRVELVSLPVNSSF 388
Qy 440 LIPS 443
Db 389 SLDS 392

RESULT 4
SNF1 CANAL
ID SNF1 CANAL STANDARD; PRT; 620 AA.
AC P52497; Q00309;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1.-).
GN SNF1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 32354 / B-311;
RX MEDLINE=98053924; PubMed=9393775;
RA Petter R., Chang Y.C., Kwon-Chung K.J.;
RT "A gene homologous to Saccharomyces cerevisiae SNF1 appears to be
RT essential for the viability of Candida albicans.";
RL Infect. Immun. 65:4909-4917 (1997).
```



```

DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_Domain.
DR Pfam; PF02149; KAl; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00300; UBA; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 42 294 PROTEIN_KINASE.
FT NP_BIND 48 56 UBA.
FT DOMAIN 315 355 UBA.
FT ACT_SITE 165 165 BY SIMILARITY.
FT BINDING 71 71 ATP (BY SIMILARITY).
FT MOD_RES 198 198 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT VAR5PLIC 1 23 Missing (in isoform 2).
FT SEQUENCE 535 AA; 61181 MW; 5535 AA; 61181 MW; PFC383223FD8317 CRC64;

Query Match 16.2%; Score 426.5; DB 1; Length 535;
Best Local Similarity 34.0%; Pred. No. 1.4e-20;
Matches 97; Conservative 59; Mismatches 98; Indels 31; Gaps 6;

QY 125 LNQYKLOSETGKGYGVVRLAYNESEDHRYAMKVLKLLKQYGFPRPPRGSQAAG 184
Db LPNKLGRITLIGSGFGRVKAIEAHLTGKVAIKLNRKI----- 78

QY 185 GPAQLPLRVRVQEIALLKLDHVNVLVLEVLDDPAEDNLYLFDLLRKGPMVEPCD 244
Db ---KNMEMEKVRREIKLFLPHPIIRLYEVIETPTD--IYLVMEYVNSGELFDYIVE 133

QY 245 K-PFSEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDDGHVVKIADFGVSNQPEGN 303
Db 134 KGRLOEAEARNFFQIISGVYCHRNVMVHRDLKPNLLDLSKCNVLIADFGSLNMRDG 193

QY 304 DAQLSSTAGTAPAFNAPAISSGSGSGKALDVWATGVTLVYCFYVYKCPFDIFILALHR 363
Db 194 HF-LKTSFGPNVAAPEVI--SGKLYAGPEVDVWSCGVILYALLCGTLPTDDENIPNLFK 250

QY 364 KIKNEPVVFPPEEPISEELKDLILKMDKNPETHIGVDPDKLHPW 408
Db 251 KIKGIYTLFS--HLSPGARDLIPRLVDPMDKRVITPEIRQHPW 293

RESULT 7
ID DCK1_MOUSE STANDARD; PRT; 756 AA.
AC Q3ULM8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase DCAMK1 (EC 2.7.1.1-) (Doublecortin-
DE like and CAM kinase-like 1).
DE DCAMK1 OR DCLK.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20004649; PubMed=10533048;
RA Burgess H.A., Martinec S., Reiner O.;
RT "KIAA0369, doublecortin-like kinase, is expressed during brain
RT development.";
RL J. Neurosci. Res. 58:567-575 (1999).
CC -!- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-

```

```

CC CC SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING
CC CC BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS
CC CC SYSTEM (by similarity).
CC CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CamK
CC CC subfamily.
CC CC -!- SIMILARITY: Contains 2 doublecortin domains.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AF155819; AAF26673.1; -.
CC CC HGSP; Q63450; 1A06.
CC CC MGD; MGI:1330861; DcamK1.1.
CC CC InterPro; IPR003533; DCX.
CC CC InterPro; IPR000719; Prot_kinase.
CC CC InterPro; IPR008271; Ser_Ehr_pkin_AS.
CC CC InterPro; IPR002290; Ser_thr_pkinase.
CC CC Pfam; PF03607; DCX; 2.
CC CC Pfam; PF00069; pkinase; 1.
CC CC ProDom; PD000001; Prot_kinase; 1.
CC CC SMART; SM00537; DCX; 2.
CC CC SMART; SM00220; S_TKc; 1.
CC CC PROSITE; PS00309; DC; 2.
CC CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC CC Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
CC CC Neurogenesis; 57 143 DOUBLECORTIN 1.
CC CC FT DOMAIN 186 269 DOUBLECORTIN 2.
CC CC FT DOMAIN 298 358 SER/PRO-RICH.
CC CC FT DOMAIN 406 663 PROTEIN_KINASE.
CC CC FT NP_BIND 412 420 ATP (BY SIMILARITY).
CC CC FT BINDING 435 435 ATP (BY SIMILARITY).
CC CC FT ACT_SITE 527 527 BY SIMILARITY.
CC CC SQ SEQUENCE 756 AA; 84153 MW; 3D1DBF18C23129F2 CRC64;

Query Match 15.7%; Score 412.5; DB 1; Length 756;
Best Local Similarity 27.6%; Pred. No. 1.8e-19;
Matches 140; Conservative 71; Mismatches 199; Indels 97; Gaps 20;

QY 35 PEPTNRGVDP-----PPRARAASVTPGSTSRLLPAPSL-SARKLSLQE 77
Db 299 PGPSRRSKSPASTSVNGTSPGSQLSTPRSGKSPSPSTSPGSLRKQRIHQSGSSTLS 358

QY 78 RPAGSYLEAQAQPYATGPAS-----HISPRWRPTIESHHVAISDAEDCVL-----NQ 127
Db 359 TKVCSSMDNDGP---GEGDELGRHSLQGMRR-----ESEEQFIPATITER 405

QY 128 YKLQSEIGKGYGVVRLAYNESEDHRYAMKVLKLLKQYGFPRPPRGSQAAGCPA 187
Db 406 YKVGRTIGDGNFAVVEKCIESTAREYALKIISKSC-----RGKE----- 446

QY 188 KQLPLERVQETAILKLDHVNVLVLEVLDDPAEDNLYLFDLLRKGPMVE-VPCKDP 246
Db 447 -----HMIQNEVSILRRVKHPNIVILLIEMDVPTD--LYLVNMLVKGGLFDLITSTK 498

QY 247 FSEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDD--DG--HVKTADFGVSNQFEG 302
Db 499 YTERDASGLMYNLASAIVKILHSLNIVHRDIKPNLLVYEHQDQSKSLKGLDGLATVDG 558

QY 303 NDAQLSSTAGTAPAFNAPAISSGSGSGKALDVWATGVTLVYCFYVYKCPFP--IDDFILA 360
Db 559 ---PLYTVCGTPTVYVAPEIIAETGY---GLKWDIWAAGVITYILLCGFPFRSGDDQEV 612

QY 361 LHRKIKNEPVVPEE--PPISEELKDLILKMDKNPETHIGVDPDKLHPWTKNGEPLP 418
Db 613 LFDQILMGQVDFPSPYWDNVDNSAKELINMMLLVNVVDQRFSAVQVLEHPWVNDG---LP 669

```



```

GIN4_YEAST          STANDARD;          PRT; 1142 AA.
ID  GIN4_YEAST
AC  Q12263;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Serine/threonine-protein kinase GIM4 (EC 2.7.1.-).
GN  GIM4 OR YDR507C OR D9719.13.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99030835; PubMed=9813093;
RA  Longtine M.S., Fares H., Pringle J.R.;
RT  "role of the yeast Gln4p protein kinase in septin assembly and the
RT  relationship between septin assembly and septin function.";
RL  J. Cell Biol. 143:719-736 (1998).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA  Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA  Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA  Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA  Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA  Winant A., Yelton M., Botstein D., Davis R.W.;
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: May play a role in septin assembly.
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC  subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U33140; AAA75513.1; -.
CC  EMBL; U33057; AAB64949.1; -.
CC  PIR; S59359; S59359.
CC  HSP; Q63450; 1A06.
CC  Germonline; 140399; -.
CC  SGD; S0002915; GIN4.
CC  GO; GO:0005935; C:bud neck; IDA.
CC  GO; GO:0007117; P:bud growth; IGI.
CC  GO; GO:000135; P:septin checkpoint; IGI.
CC  GO; GO:0000921; P:septin ring assembly; IGI.
CC  InterPro; IPR000719; Prot_kinase.
CC  InterPro; IPR008271; Ser_thr_pkin_AS.
CC  InterPro; IPR002290; Ser_thr_pkinase.
CC  Pfam; PF00069; pkinase; 1.
CC  PRINTS; PR00109; TYRKINASE.
CC  ProDom; PD000001; Prot_kinase; 1.
CC  SMART; SMW0220; S_TKC; 1.
CC  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC  PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC  PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC  Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC  DOMAIN 19 289 PROTEIN KINASE.
CC  NP_BIND 25 33 ATP (BY SIMILARITY).
CC  BINDING 48 48 ATP (BY SIMILARITY).
CC  ACT_SITE 156 156 BY SIMILARITY.
CC  SEQUENCE 1142 AA; 129857 MW; EC16FF4BB49DD811 CRC64;
Query Match 15.3%; Score 403; DB 1; Length 1142;
Best Local Similarity 32.4%; Pred. No. 1.2e-18;
Matches 119; Conservative 57; Mismatches 143; Indels 48; Gaps 12;
QY 125 LNQYKLAQSEIGKGVGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQAQG 184

```

```

Db 16 IGFWKLGELGLGSGTKVQLARNGSTGQEAARVISKAVF-----NTGVSGTIVG 67
QY 185 GPAKQLLP--LERYQEIAILKKLDHVNVLKLEVLDDPAEDNLVLFVDFLLRKGPNVEVP 242
Db 68 STTPDALPYGIER---EIIIMKLLNHPNVLRLDYDWE--TNTDLVLYLVEAERGFNL 122
QY 243 CDK-PFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDDGHVKTADFGVSQF 301
Db 123 VERGPLPEHEAIRFFQIIGVSYCHALGIVHRDLKPNLLOHKNIKIADFGWA-ALE 181
QY 302 GNDALQSSTAGTAPAFMAPRAISDGSFGSKALDVWATGVTLYCFVYVGKCFP--IDDFIL 359
Db 182 TECKLLETSCGSPHYAAPEIV--SGIPYQGFASDVMSGVLFALLTGRLPFDEEDGNIR 239
QY 360 ALHRKIKNEVPVFPPEPEISEELKDLILKMLDKNPETRICVPDIKHPWTK-----N 412
Db 240 TLLLVKQGFEMPSDDEISREAQDLIRKILTVDPERRIKTRDILKHPLLQKYPISIRDSK 299
QY 413 GEFPDLPSEERHSCVVEVTEGEVKNVRLIPSWTIVL-----VKSMRKRSFGNPFE 464
Db 300 SINGLPREDIYLTPL-----SESNSIDATILQNLVILWHGRDPEGIKEKL-R-----E 347
QY 465 PQARREE 471
Db 348 PGANAEX 354
RESULT 13
DCK1 RAT          STANDARD;          PRT; 433 AA.
ID  DCK1 RAT
AC  O08875;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-
DE  like and CAM kinase-like 1) (Calcium/calmodulin-dependent protein
DE  kinase type I-like CPG16).
DE  DCAMKL1 OR CPG16.
GN  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=101116;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Wistar;
RX  MEDLINE=98364306; PubMed=9699150;
RA  Hevroni D., Rattner A., Bundman M., Lederfein D., Gabard A.,
RA  Mangelus M., Silverman M.A., Kedar H., Naor C., Kornuc M., Hanoch T.,
RA  Seger R., Theill L.E., Nedivi E., Richter-Levin G., Citri Y.;
RA  "Hippocampal plasticity involves extensive gene induction and multiple
RA  cellular mechanisms.";
RT  J. Mol. Neurosci. 10:75-98 (1998).
RL  -!- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-
CC  SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING
CC  BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS
CC  SYSTEM (BY SIMILARITY).
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CamK
CC  subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U78857; AAC99476.1; -.
CC  HSP; Q63450; 1A06.
CC  InterPro; IPR000719; Prot_kinase.
CC  InterPro; IPR008271; Ser_thr_pkin_AS.
CC  InterPro; IPR002290; Ser_thr_pkinase.

```



```
DR pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Neurogenesis.
FT DOMAIN 1 33 SER/PRO-RICH (BY SIMILARITY) .
FT DOMAIN 83 340 PROTEIN KINASE.
FT DOMAIN 391 394 POLY-ARG.
FT NP_BIND 89 97 ATP (BY SIMILARITY).
FT BINDING 112 112 ATP (BY SIMILARITY).
FT ACT_SITE 204 204 BY SIMILARITY.
FT ACT_SITE 204 204 BY SIMILARITY.
SQ SEQUENCE 433 AA; 47680 MW; 0CE5E06E152A557D CRC64;

Query Match 15.3%; Score 402.5; DB 1; Length 433;
Best Local Similarity 27.9%; Pred. No. 3.9e-19;
Matches 128; Conservative 69; Mismatches 179; Indels 83; Gaps 17;

QY 81 GSYLEAQAGPYATGPASHISPAWRRPTTIESH---HVAISDAEDCVQLNQ----- 127
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 12 GSQSTPSRGKSPSP-SPTSPGSLRKQRTSQHGSSTLSSTKVCSSMDENDGPGEESD 70
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 128 -----YKQSEICKGAYGVVRLAYNESEDRHYAMKVLKSKLLKQYGPFRPP 175
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 71 EGFQIPATITERKYKVRTIGDGNFAVVKECIERSTAREYALKIIRKSKC----- 119
QY 176 PRGSAQAQGPAPKQLPLERVVQETAILKKLDHVVVVKLIEVLDDPAEDNLVLPDLRK 235
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 120 -RGKE-----HMQNEVSILRRVXHPNVLIIIEMDVPE--LYLWELVKG 163
QY 236 GVWME-VPCKPFBSEARLYLRDLVILGLEYLHCQKIVHRDIKPSNLLGD--DG--HYK 290
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 164 GDLFAITSTSKYTERDASGLMYLNLSAIAKYLHSLNIVHRDIKPNLIVYHQDGSKSK 223
QY 291 INDFGVNQFEGNDQALSTAGTAPFMAPEALSDSQSGKALDWATGVTLYCFVYCK 350
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 224 LGDFGLATVDSG---PLYTVCGTPTPYVAPEIIAETGY---GLKVDIWAAGVITYLLCGF 277
QY 351 CFPF--IDDFILALHRKIKNEPVVFPPEE--PETSEELKOLILKMLDKNPETRIGVDPDKLH 406
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 278 PFRSGSDQDQVLFQILMGQVDFSPYDNDVSDSAKELINMLLVNDQRFSAVQVLEH 337
QY 407 PWTXKNGEELPSEBEHCSWVEVTRGEVKNVRLIPSTVTVILVKSMLKRSFGNPFEPQ 466
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 338 PWNVDG--LPENEHQLSVA---GKIKHPNTGPKSPSTAAGSVVIATTALDK--ERQ 388
QY 467 ARREERSMSAPGNLLVKGFGGKGKSPGLPGVQDEAAS 505
Dd ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd 389 VERRRRNQDVRGRYKAQPA-----PPELNSSESDYSPS 421

RESULT 14
DCK1_HUMAN
ID DCK1_HUMAN STANDARD; PRT; 740 AA.
AC O15075;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase DCAMK1 (EC 2.7.1.-) (Doublecortin-
DE like and CM kinase-like 1).
GN DCAMK1 OR KIAA0369.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RA
```

```
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fetal brain;
RX MEDLINE=99156863; PubMed=10036192;
RA Sossey-Alaoui K., Srivastava A.K.;
RT "DCAMK1, a brain-specific transmembrane protein on 13q12.3 that is
RT similar to doublecortin (DCX).";
RL Genomics 56:121-126(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
RX MEDLINE=98419166; PubMed=9747029;
RA Omori Y., Suzuki M., Ozaki K., Harada Y., Nakamura Y.,
RA Takahashi E.-I., Fujiwara T.;
RT "Expression and chromosomal localization of KIAA0369, a putative
RT kinase structurally related to Doublecortin.";
RL J. Hum. Genet. 43:169-177(1998).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=99162404; PubMed=10051403;
RA Matsumoto N., Pilz D.F., Ledbetter D.H.;
RT "Genomic structure, chromosomal mapping, and expression pattern of
RT human DCAMK1 (KIAA0369), a homologue of DCX (XLIS).";
RL Genomics 56:179-183(1999).
CC -!- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-
CC SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING
CC BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS
CC SYSTEM.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Comment-Additional isoforms seem to exist. Type A (AS and AL)
CC and type B (BS and BL) isoforms differ respectively by the
CC presence or absence of the doublecortin domain. An alternative
CC splicing occurring in 3' of the mRNA produces the long (L)
CC instead of the short (S) isoforms;
CC Name=2; Synonyms=AL;
CC IsoId=O15075-1; Sequence=Displayed;
CC Name=1; Synonyms=AS;
CC IsoId=O15075-2; Sequence=VSP_004907;
CC Name=3; Synonyms=BS;
CC IsoId=O15075-3; Sequence=VSP_004905, VSP_004906, VSP_004907;
CC Name=4; Synonyms=BL;
CC IsoId=O15075-4; Sequence=VSP_004905, VSP_004906;
CC -!- TISSUE SPECIFICITY: IN FETAL TISSUES, HIGHLY EXPRESSED IN BRAIN,
CC DETECTABLE IN LUNG AND LIVER, BUT NOT IN KIDNEY. IN ADULT TISSUES,
CC EXPRESSED UBQUITOUSLY IN THE BRAIN, DETECTABLE IN THE HEART,
CC LIVER, SPLEEN, THYMUS, PROSTATE, TESTIS, OVARY, SMALL INTESTINE
CC AND COLON. THE TYPE A ISOFORMS SEEM TO BE EXPRESSED PREDOMINANTLY
CC IN FETAL BRAIN WHEREAS TYPE B ISOFORMS ARE EXPRESSED ABUNDANTLY IN
CC BOTH FETAL AND ADULT BRAIN.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CamK
CC subfamily.
CC -!- SIMILARITY: Contains 2 doublecortin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AB002367; BAA20824.1; -.
CC HSP; Q63450; IA06.
CC GENE; HGNC:2700; DCAMK1.
CC MIM; 604742; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004672; F:protein kinase activity; TAS.
CC GO; GO:0005057; F:receptor signaling protein activity; TAS.
CC GO; GO:0007417; P:central nervous system development; TAS.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 01:17:43 ; Search time 102 Seconds
(without alignments)

1398.887 Million cell updates/sec

Title: US-10-690-617-2

Perfect score: 2834
Sequence: 1 MEGAPAYCCDPRAELRV.....FGEKGKSPFLGVQDEAAS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2634	100.0	505	5	ABB83054 Human kin
2	2626	99.7	505	5	AAE19161 Human kin
3	2626	99.7	505	7	ADE61939 Human pro
4	2626	99.7	505	7	ADE61943 Human pro
5	2619	99.4	582	5	ABP69334 Human pol
6	2611	99.1	505	5	ABU65068 Human NOV
7	2595	98.5	513	4	AAU03510 Human pro
8	2595	98.5	513	4	AAE04361 Human kin
9	2591	98.4	561	5	AAU79458 Human nov
10	2585	98.1	543	5	ABU65069 Human NOV
11	2560	97.2	503	5	ABU65067 Human NOV
12	2468	93.7	505	5	ABB83055 Rat amino
13	2468	93.7	505	7	ADE61937 Rat Prote
14	2468	93.7	505	7	ADE61941 Rat Prote
15	1537.5	58.4	540	3	AAV44239 Human cel
16	1525.5	57.9	556	7	ADB75228 Prostate
17	1509	57.3	680	4	AAU40450 Human pol
18	1509	57.3	680	4	AAU40449 Human pol
19	1504	57.1	588	4	AAU38663 Human pol
20	1498	56.9	588	4	AAU38659 Novel pro
21	1472.5	55.9	417	5	ABG61855 Prostate
22	1472.5	55.9	417	7	ADB75230 Prostate
23	1342	50.9	498	7	AUC99093 Human RPP
24	1308.5	49.7	545	4	AAU38664 Human pol
25	923	35.0	210	4	AAU25244 Human pro

26	881	33.4	338	4	ABB66416	Abb66416 Drosophil
27	648	24.6	423	6	ABU11714	Abu11714 Human MDD
28	611	23.2	140	4	AAU25471	AAU25471 Human pro
29	566	21.5	1142	6	ABR52943	ABR52943 Protein s
30	526	20.0	250	5	ABB78798	Abb78798 Serine/th
31	452	17.2	570	6	ABR40710	ABR40710 Zea may9
32	436.5	16.6	511	2	AAU40842	AAU40842 SHPP. 8/2
33	430.5	16.3	523	3	AAU03425	AAU03425 Wheat put
34	430.5	16.3	523	6	ABR40719	ABR40719 Triticum
35	429.5	16.3	422	3	AAU03419	AAU03419 Soybean p
36	429.5	16.3	422	6	ABR40713	ABR40713 Glycine m
37	428.5	16.3	438	3	AAU03421	AAU03421 Soybean p
38	428.5	16.3	438	6	ABR40715	ABR40715 Glycine m
39	428	16.2	441	3	AAU03423	AAU03423 Soybean p
40	428	16.2	441	6	ABR40717	ABR40717 Glycine m
41	427	16.2	254	5	ABB78797	Abb78797 Protein k
42	426.5	16.2	512	6	ABR44015	ABR44015 Human SNF
43	426.5	16.2	1349	4	AAU70854	AAU70854 C albican
44	424.5	16.1	579	3	AAU03417	AAU03417 Corn puta
45	424.5	16.1	579	6	ABR40709	ABR40709 Zea may5

ALIGNMENTS

RESULT 1					
ABB83054					
ID	ABB83054	standard; protein; 505 AA.			
XX	AC	ABB83054;			
XX	DT	27-AUG-2002 (first entry)			
XX	DE	Human kinase amino acid sequence.			
XX	KW	Human; kinase protein; calcium/calmodulin-dependant protein kinase; eye;			
XX	KW	retinoblastoma; brain; kinase modulator.			
XX	OS	Homo sapiens.			
XX	FT	Key	Location/Qualifiers		
FT	Region	26..29	/label= casein_kinase_II_phosphorylation_site		
FT	Region	58..60	/label= protein_kinase_C_phosphorylation_site		
FT	Region	69..71	/label= protein_kinase_C_phosphorylation_site		
FT	Region	71..74	/note= "cAMP- and cGMP-dependant protein kinase phosphorylation site"		
FT	Region	74..77	/label= casein_kinase_II_phosphorylation_site		
FT	Region	82..85	/label= casein_kinase_II_phosphorylation_site		
FT	Region	100..102	/label= protein_kinase_C_phosphorylation_site		
FT	Region	105..108	/note= "cAMP- and cGMP-dependant protein kinase phosphorylation site"		
FT	Region	117..120	/label= casein_kinase_II_phosphorylation_site		
FT	Region	134..157	/note= "protein kinase ATP-binding region signature"		
FT	Region	147..150	/label= N-glycosylation_site		
FT	Region	160..162	/label= protein_kinase_C_phosphorylation_site		
FT	Region	178..183	/label= N-myristoylation site		
FT	Region	271..283	/note= "serine/threonine protein kinase active-site signature"		
FT	Region	326..331			

FT Region /label= N-myristoylation site O
 FT 330..332
 FT /label= protein_kinase_C_phosphorylation_site
 FT Misc-difference 375
 FT /note= "wild-type Glu is replaced by Gly as a result of
 FT an SNP (single nucleotide polymorphism) in the encoding
 FT DNA"
 FT 419..422
 FT /label= casein_kinase_II_phosphorylation_site
 FT 425..428
 FT /label= casein_kinase_II_phosphorylation_site
 FT 430..433
 FT /label= casein_kinase_II_phosphorylation_site
 FT 437..439
 FT /label= protein_kinase_C_phosphorylation_site
 FT 455..458
 FT /note= "cAMP- and cGMP-dependant protein kinase
 FT phosphorylation site"
 FT 485..492
 FT /label= ATP/GTP-binding_site_motif_A
 FT /note= "P-loop"
 FT
 FT WO200224920-A2.
 FT
 FT 28-MAR-2002.
 FT
 FT 19-SEP-2001; 2001WO-US029161.
 FT
 FT 19-SEP-2000; 2000US-0233493P.
 FT 13-NOV-2000; 2000US-0247031P.
 FT 06-DEC-2000; 2000US-0072995.
 FT
 FT (PEKE) PE CORP NY.
 FT
 FT Beasley EM, Wei M, Bonazzi VR, Sanders R, Di Francesco V;
 FT WPI; 2002-404955/43.
 FT N-PSDB; ABL58698.
 FT
 FT Novel peptide designated as human kinase useful as target for diagnosing
 FT a disease or predisposition to the disease mediated by the peptide.
 FT
 FT Claim 1a; Fig 2; 89pp; English.
 FT
 FT The invention relates to an isolated peptide designated human kinase
 FT (HK) that has homology to members of the calcium/calmodulin-dependent
 FT protein kinase kinase subfamily. The mechanism of action of the protein
 FT of the invention is that of a kinase modulator. The human kinase of the
 FT invention is useful for creating a pharmaceutical composition for
 FT treating a disease or condition mediated by the human kinase. HK is also
 FT useful to provide a target for diagnosing a disease or predisposition to
 FT disease mediated by HK, and is also useful in pharmacogenic analysis. HK
 FT is useful for treating a disorder characterised by absence of
 FT inappropriate or unwanted expression of HK, also as an immunogen to raise
 FT antibodies by administering HK to a mammalian organism e.g. rat, rabbit
 FT or mouse. Nucleic acids of the invention are useful as hybridisation
 FT probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells
 FT expressing a native HK are useful for assaying compounds that stimulate
 FT or inhibit HK function. Nucleic acids of the invention are also useful
 FT for producing transgenic animals. Experimental data indicates that kinase
 FT proteins of the present invention are expressed in humans in the eye
 FT (retinoblastomas) and brain. The current sequence represents the amino
 FT acid sequence of the human kinase of the invention
 FT
 FT Sequence 505 AA;
 FT
 FT Query Match 100.0%; Score 2634; DB 5; Length 505;
 FT Best Local Similarity 100.0%; Pred. No. 2.8e-236;
 FT Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 FT
 FT 1 MEGGPAVCCDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 FT
 FT 1 MEGGPAVCCDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 FT

QY 61 LLPARSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRARRPTTIESHHVAISDAE 120
 DB 61 LLPARSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRARRPTTIESHHVAISDAE 120
 QY 121 DCVOLNQYKLSQSEIGKAGVVRLAYNESDRHYAMKVLSSKKLLKQYGFPRPPRRGSQ 180
 DB 121 DCVOLNQYKLSQSEIGKAGVVRLAYNESDRHYAMKVLSSKKLLKQYGFPRPPRRGSQ 180
 QY 181 AAGGPAKQLPLERVYQEIALLKLDHVNKKLIEVLDDPAEDNLYLVFDLLKRGPMVE 240
 DB 181 AAGGPAKQLPLERVYQEIALLKLDHVNKKLIEVLDDPAEDNLYLVFDLLKRGPMVE 240
 QY 241 VPCDKPFSBEQARLYLRDVILGLEYLHCQKIVHRDIKPSNLLIGDDGHVKIADPGVSNQF 300
 DB 241 VPCDKPFSBEQARLYLRDVILGLEYLHCQKIVHRDIKPSNLLIGDDGHVKIADPGVSNQF 300
 QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVWATGVTLYCFVYGKCPIDDFILA 360
 DB 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVWATGVTLYCFVYGKCPIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNPETRIQVGPDIKLPHPWTKNGEELPSE 420
 DB 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNPETRIQVGPDIKLPHPWTKNGEELPSE 420
 QY 421 EEHCSVVEVTEGEVKNVRLIPSWTTVILVKMLRKRSFGNPPFPQARRERSMSAPGNL 480
 DB 421 EEHCSVVEVTEGEVKNVRLIPSWTTVILVKMLRKRSFGNPPFPQARRERSMSAPGNL 480
 QY 481 LVKEGFGEGGKSPELPGVQDEAAS 505
 DB 481 LVKEGFGEGGKSPELPGVQDEAAS 505
 RESULT 2
 ID AAE19161 standard; protein; 505 AA.
 XX
 AC AAE19161;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human kinase polypeptide (PKIN-19).
 XX
 KW Human; kinase polypeptide; PKIN-19; gene therapy; Addison's disease;
 KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
 KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
 KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
 KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
 KW cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease;
 KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
 KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
 KW hypotensive; anti-HIV; enzyme.
 XX
 KW Homo sapiens.
 XX
 OS
 FH Key Location/Qualifiers
 FT Domain 128..409
 FT Binding-site /note= "Eukaryotic proetin kinase domain"
 FT 485..492
 FT /note= "ATP/GTP-binding site motif A"
 XX
 PN WO200208399-A2.
 XX
 PD 31-JAN-2002.
 XX
 XX 20-JUL-2001; 2001WO-US023092.
 XX
 XX 21-JUL-2000; 2000US-0220038P.
 PR 28-JUL-2000; 2000US-0222112P.
 PR 04-AUG-2000; 2000US-0222831P.
 PR 11-AUG-2000; 2000US-0224729P.
 XX

PA (INCY-) INCYTE GENOMICS INC.
PA (THOR/) THORNTON M.
XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK,
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;
PI Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;
PI Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;
XX WPI; 2002-206083/26.
DR N-PSDB; AAD30566.
XX
PT New human kinase polypeptide, useful in diagnosis, prevention and
PT treatment of cancer, immune disorder, growth and developmental disorder,
PT cardiovascular disorder and lipid disorder.
XX
PS Claim 1; Page 168-170; 196pp; English.
XX
CC The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
CC present sequence is human PKIN-19
XX
SQ Sequence 505 AA;
Query Match 99.7%; Score 2626; DB 5; Length 505;
Best Local Similarity 99.8%; Pred. No. 1.6e-235;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEGGPVAVCCQPRAEIVRVAIDVTHLEADGGPEPTRNGVDPVPPRARAASVTPGSTSR 60
DB 1 MEGGPVAVCCQPRAEIVRVAIDVTHLEADGGPEPTRNGVDPVPPRARAASVTPGSTSR 60
QY 61 LLPAPSLARKLSIQRPAGSYLEAQAQGYATGPAGHSIPRWRRTTIESHHVAISDAE 120
DB 61 LLPAPSLARKLSIQRPAGSYLEAQAQGYATGPAGHSIPRWRRTTIESHHVAISDAE 120
QY 121 DCVQLNQYKLSQKSGKAYGVVRLAYNESDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
DB 121 DCVQLNQYKLSQKSGKAYGVVRLAYNESDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
QY 181 AAQGGPAKQLPLERVQETIAILKLDHVNIVKLVLEVLDDPAEDNLVLFLLRGPVME 240
DB 181 AAQGGPAKQLPLERVQETIAILKLDHVNIVKLVLEVLDDPAEDNLVLFLLRGPVME 240
QY 241 VPCDKPSEQARLYRDLVILGLEYLHCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
DB 241 VPCDKPSEQARLYRDLVILGLEYLHCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPFAPEAIDSQGSFGKALDWMATGVTLYCFVYKCPFFIDDFILA 360
DB 301 EGNDAQLSSTAGTAPFAPEAIDSQGSFGKALDWMATGVTLYCFVYKCPFFIDDFILA 360
QY 361 LHRKIKNEPVVFPPEPEISEBELKDLILKMLDKNPETRIGVDDIKLHPWVTNKGEEPLPSE 420
DB 361 LHRKIKNEPVVFPPEPEISEBELKDLILKMLDKNPETRIGVDDIKLHPWVTNKGEEPLPSE 420
QY 421 EEHCSSVVEVEGEVKNVRLIPSWTTTIVLVKSMRLKESFGNPFEPQARRERSMSAPCNL 480
DB 421 EEHCSSVVEVEGEVKNVRLIPSWTTTIVLVKSMRLKESFGNPFEPQARRERSMSAPCNL 480
QY 481 LVXEGFGGKGKSPGLPGVQDEAAS 505
DB 481 LVXEGFGGKGKSPGLPGVQDEAAS 505
RESULT 3
ADE61939
ID ADE61939 standard; protein; 505 AA.
XX
AC ADE61939;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_115670, SEQ ID NO 7868.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
FN WO2003016475-A2.
XX
ED 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; NP_115670.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
CC
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 505 AA;
 Query Match 99.7%; Score 2626; DB 7; Length 505;
 Best Local Similarity 99.8%; Pred. No. 1.6e-235;
 Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEGGPVCCODPRAELVERVAADVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 DB 1 MEGGPVCCODPRAELVERVAADVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 QY 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPPTTIESHVAISDAE 120
 DB 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPPTTIESHVAISDAE 120
 QY 121 DCVQLNQYKLOSEIGKAGYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPPRPPRGSQ 180
 DB 121 DCVQLNQYKLOSEIGKAGYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPPRPPRGSQ 180
 QY 181 AAQGGPAKQLPLERVYQETAILKKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
 DB 181 AAQGGPAKQLPLERVYQETAILKKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
 QY 241 VPCDKPFSEEQARLYLRDVLIGLYHCOKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
 DB 241 VPCDKPFSEEQARLYLRDVLIGLYHCOKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
 QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVMATGVTLYCFYVYKCPFIDDFILA 360
 DB 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVMATGVTLYCFYVYKCPFIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKHPWTKNGEELPSE 420
 DB 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKHPWTKNGEELPSE 420
 QY 421 EEHCVVVEVTEGEVKNVRLIPSWTIVLVKSMRKRSFGNPPPEQARREERSMSAPGNL 480
 DB 421 EEHCVVVEVTEGEVKNVRLIPSWTIVLVKSMRKRSFGNPPPEQARREERSMSAPGNL 480
 QY 481 LVKEGFGEGGKSPELPQVQDEAAS 505
 DB 481 LVKEGFGEGGKSPELPQVQDEAAS 505

RESULT 4
 ID ADE61943
 XX ADE61943 standard; protein; 505 AA.
 AC ADE61943;
 XX XX
 DT 29-JAN-2004 (first entry)
 XX Human Protein NP_115670, SEQ ID NO 7872.
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS
 XX WC2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR
 XX 01-NOV-2001; 2001US-0346362P.
 PR
 XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK, NP_115670.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 505 AA;

Query Match 99.7%; Score 2626; DB 7; Length 505;
 Best Local Similarity 99.8%; Pred. No. 1.6e-235;
 Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEGGPVCCODPRAELVERVAADVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 DB 1 MEGGPVCCODPRAELVERVAADVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 QY 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPPTTIESHVAISDAE 120
 DB 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPPTTIESHVAISDAE 120
 QY 121 DCVQLNQYKLOSEIGKAGYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPPRPPRGSQ 180
 DB 121 DCVQLNQYKLOSEIGKAGYGVVRLAYNESEDHRYAMKVLKSKLLKQYGFPPRPPRGSQ 180
 QY 181 AAQGGPAKQLPLERVYQETAILKKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
 DB 181 AAQGGPAKQLPLERVYQETAILKKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
 QY 241 VPCDKPFSEEQARLYLRDVLIGLYHCOKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
 DB 241 VPCDKPFSEEQARLYLRDVLIGLYHCOKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
 QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVMATGVTLYCFYVYKCPFIDDFILA 360
 DB 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSFSGKALDVMATGVTLYCFYVYKCPFIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKHPWTKNGEELPSE 420

PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 18-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DW, Rieger DK;
XX
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97035.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.

XX Claim 1; Page 139; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 505 AA;

Query Match 99.1%; Score 2611; DB 5; Length 505;
Best Local Similarity 99.4%; Pred. No. 3.9e-234;
Matches 502; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEGGPAVCCDPAELVERVAALDVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
DB 1 MEGGPAVCCDPAELVERVAALDVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
QY 61 LLPAREPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAWRPTTIESHHVAISDAE 120
DB 61 LLPAREPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAWRPTTIESHHVAISDAE 120
QY 121 DCVQLNQYKLSQSEIGKAGYGVRLAYNESEDHRYAMKVLKKKLLKQYGFPPRPPRGSQ 180
DB 121 DCVQLNQYKLSQSEIGKAGYGVRLAYNESEDHRYAMKVLKKKLLKQYGFPPRPPRGSQ 180
QY 181 AAOGGPAKOLLPLERYQEIALLKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
DB 181 AAOGGPAKOLLPLERYQEIALLKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
QY 241 VPCDKPFSEEQARLYLRDVLGLYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
DB 241 VPCDKSFSEEQARLYLRDVLGLYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPMAPEAISDSGSGKALDYWATGVTLYCFYVGCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPMAPEAISDSGSGKALDYWATGVTLYCFYVGCPIDDFILA 360
QY 361 LHRKIKNEPVWPEPEI SEELKDILKMLDKNPNETRIGVPDIKLPWTKNGEELPSE 420
DB 361 LHRKIKNEPVWPEPEI SEELKDILKMLDKNPNETRIGVPDIKLPWTKNGEELPSE 420
QY 421 EEHCSSVVEVTEGVKNSVRLIPSWTTVLVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
DB 421 EEHCSSVVEVTEGVKNSVRLIPSWTTVLVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
QY 481 LVKEGFGEGGKSPQLPGVQDEAAS 505
DB 481 LVKEGFGEGGKSPQLPGVQDEAAS 505

RESULT 7
AAU03510
ID AAU03510 standard; protein; 513 AA.
XX
XX AAU03510;
XX
DT 12-SEP-2001 (first entry)
XX
XX Human protein kinase #10.
DE
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX

PN WO200138503-A2.
XX 31-MAY-2001.
PD 22-NOV-2000; 2000WO-US032085.
XX 24-NOV-1999; 99US-0167482P.
PR (SUGEN-) SUGEN INC.
PA Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI: 2001-343950/36.
DR N-PSDB; AAS06710.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Claim 7; Fig 2; 433pp; English.
XX
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 513 AA;
Query Match 98.5%; Score 2595; DB 4; Length 513;
Best Local Similarity 97.9%; Pred. No. 1.3e-232;
Matches 503; Conservative 0; Mismatches 1; Indels 10; Gaps 2;
Qy 1 MEGGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTRNGVDPPRARAASVIPGSTSR 60
Db 1 MEGGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTRNGVDPPRARAASVIPGSTSR 60
Qy 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPAWRPTIESHHVAISDAE 120
Db 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPAWRPTIESHHVAISDAE 120
Qy 121 DCVQLNOYKLSQSEIGK-----GAYGVVRLAYNESEDRHYAMKVLKSKKLKQYGF 171
Db 121 DCVQLNQYKLSQSEIGKVGLTDAYLQAGYGVVRLAYNESEDRHYAMKVLKSKKLKQYGF 180
Qy 172 RRPFRGSQAAGGPAKQLLPLERYQIEAILKKLDHNVVVKLIEVLDDPAEDNLVLFVD 231
Db 181 RRPFRGSQAAGGPAKQLLPLERYQIEAILKKLDHNVVVKLIEVLDDPAEDNLVLFV-D 239
Qy 232 LLRKGPMVEVCDKPFSEQARLYLRDVLGLYLHCOKIVHRDIKPSNLLLGDDGHVKI 291
Db 240 LLRKGPMVEVCDKPFSEQARLYLRDVLGLYLHCOKIVHRDIKPSNLLLGDDGHVKI 299
Qy 292 ADFGVSNQFEGNDQALSSTAGTAPFAPEATSDSQSGSKALDVMATGVTLYCFVYKGC 351
Db 300 ADFGVSNQFEGNDQALSSTAGTAPFAPEATSDSQSGSKALDVMATGVTLYCFVYKGC 359
Qy 352 PFIDDFILALHRKIKNEPVVFPFEPSEELKDILKMLDKNPETRIGVPDIKLPWYTK 411
Db 360 PFIDDFILALHRKIKNEPVVFPFEPSEELKDILKMLDKNPETRIGVPDIKLPWYTK 419

Qy 412 NGEPLPSEEEHCSSVVEVTEGEVQKNSVRLIPSWTTVILVKSLMKRKSFGNPFEPQARREE 471
Db 420 NGEPLPSEEEHCSSVVEVTEGEVQKNSVRLIPSWTTVILVKSLMKRKSFGNPFEPQARREE 479
Qy 472 RMSAPGNLLVKEGFGEGGKSPGLPGVQDEAAS 505
Db 480 RMSAPGNLLVKEGFGEGGKSPGLPGVQDEAAS 513

RESULT 8
AAE04361
ID AAE04361 standard; protein; 513 AA.
XX AC AAE04361;
XX DT 04-SEP-2001 (first entry)
XX DE Human kinase (PKIN)-2.
XX KW Human kinase; PKIN-2; therapy; immune disorder; Addison's disease; AIDS;
KW acquired immune deficiency syndrome; growth and developmental disorder;
KW arteriosclerosis; mixed connective tissue disease; MCTD; adenocarcinoma;
KW leukaemia; cardiovascular disease; myocardial infarction; hypertension;
KW lipid disorder; cancer; fatty liver; cholestasis; transgenic animal;
KW gene therapy; antiallergic; antiasthmatic; antithyroid; dermatological;
KW antidiabetic; nephrotrophic; antiulcer; antiarthritic; antirheumatic;
KW antipsoriatic; neuroprotective; cytostatic; hepatotrophic; osteopathic;
KW vasotropic; antianginal; anorectic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX Region 1..127
FT /note= "Kinase protein beta"
FT Domain 130..408
FT /note= "Protein kinase domain"
FT Domain 145..417
FT /note= "Eukaryotic protein kinase domain"
FT Domain 273..291
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 320..330
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 342..364
FT /note= "Tyrosine kinase catalytic domain"
FT Binding-site 493..500
FT /note= "ATP/GTP-binding site motif A (P-loop)"
XX WO200146397-A2.
XX PD 28-JUN-2001.
XX PF 20-DEC-2000; 2000WO-US035304.
XX PR 23-DEC-1999; 99US-0172066P.
PR 14-JAN-2000; 2000US-0176107P.
PR 21-JAN-2000; 2000US-017731P.
PR 28-JAN-2000; 2000US-0178573P.
XX (INCY-) INCYTE GENOMICS INC.
XX PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI Yue H, Yao MG, Lal P, Khan FA;
XX WPI: 2001-418059/44.
DR N-PSDB; AAD08635.
XX
PT Novel human kinase proteins (PKIN) useful for diagnosing, treating,
PT preventing immune disorders, cardiovascular diseases and disorders
PT affecting growth and development associated with abnormal expression of
PT PKIN.
XX
PS Claim 1; Page 104-105; 128pp; English.

Db 1 MEGPAVCCQDRAELVERAAIDVTHLEADGGPEPTRNGVDDPPRARAASVTPGSTSR 60
Qy 61 LLPARPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAMRPTIESHHVAISDAE 120
Db 61 LLPARPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAMRPTIESHHVAISDAE 120
Qy 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDRHYAMKVLKKLLKQYGFRRPPRGSG 180
Db 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDRHYAMKVLKKLLKQYGFRRPPRGSG 180
Qy 181 AAQGGPAKQLPLERVYQEIALLKLDHVVNVKLIIEVLDPAEDNLVLFDLRLKGPVME 240
Db 181 AAQGGPAKQLPLERVYQEIALLKLDHVVNVKLIIEVLDPAEDNLVLFDLRLKGPVME 240
Qy 241 VPCDKPFSEEQARLYLRDVLGLVYLHCQKIYHRDIKPSNLLGDDGHVKIADFGVSNQF 300
Db 241 VPCDKPFSEEQARLYLRDVLGLVYLHCQKIYHRDIKPSNLLGDDGHVKIADFGVSNQF 300
Qy 301 EGNDAQLSSTAGTAFMAPEAISDSQSGPSFKALDVWATGVTLYCFVYKGPFFIDDFILA 360
Db 301 EGNDAQLSSTAGTAFMAPEAISDSQSGPSFKALDVWATGVTLYCFVYKGPFFIDDFILA 360
Qy 361 LHRKIKNEPVVFPPEPEISEELKDILKMLDKNPETRIGVDPDKLHPWVTXNGEPLPSE 420
Db 361 LHRKIKNEPVVFPPEPEISEELKDILKMLDKNPETRIGVDPDKLHPWVTXNGEPLPSE 420
Qy 421 EEHCSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
Db 421 EEHCSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
Qy 481 LVKEGFGGKSPELPGVQ 499
Db 481 LVKEGFGGKSPELPGVQ 499

RESULT 10
ABU65069
ID ABU65069 standard; protein; 543 AA.
XX AC ABU65069;
XX XX
XX 20-MAY-2003 (first entry)
XX Human NOV15c protein.
XX DE
XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
XX KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
XX KW human.
XX OS Homo sapiens.
XX XX
XX PN WO200272757-A2.
XX XX
XX PD 19-SEP-2002.
XX XX
XX EF 08-MAR-2002; 2002WO-US006908.
XX XX
XX 08-MAR-2001; 2001US-0274101P.
XX PR 08-MAR-2001; 2001US-0274194P.
XX PR 08-MAR-2001; 2001US-0274281P.
XX PR 08-MAR-2001; 2001US-0274322P.
XX PR 09-MAR-2001; 2001US-0274849P.
XX PR 12-MAR-2001; 2001US-0275235P.
XX PR 13-MAR-2001; 2001US-0275578P.
XX PR 13-MAR-2001; 2001US-0275579P.
XX PR 13-MAR-2001; 2001US-0275601P.
XX PR 14-MAR-2001; 2001US-0276000P.
XX PR 16-MAR-2001; 2001US-0276776P.
XX PR 19-MAR-2001; 2001US-0276994P.
XX PR 20-MAR-2001; 2001US-0277239P.
XX PR 20-MAR-2001; 2001US-0277321P.
XX PR 20-MAR-2001; 2001US-0277327P.
XX PR 21-MAR-2001; 2001US-0277791P.

PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 13-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 02-MAY-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294899P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 18-OCT-2001; 2001US-0325681P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332722P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0333272P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX XX
XX (CURA-) CURAGEN CORP.
XX XX
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
XX Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
XX Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
XX Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
XX Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
XX N-PSDB; AEX97036.
XX XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
XX a disorder associated with aberrant NOVX expression or activity e.g.,
XX cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
XX asthma.
XX Claim 1; Page 140; 1103pp; English.
XX XX
XX This invention describes novel human NOVX polypeptides which have
XX cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
XX activity. Pharmaceutical compositions comprising the NOVX proteins or
XX nucleic acid molecules or NOVX antibodies are useful for preventing or
XX treating a disorder associated with aberrant NOVX expression or activity
XX e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial

CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU5041-ABU6218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185

XX SQ Sequence 543 AA;

Query Match 98.1%; Score 2585; DB 5; Length 543;
Best Local Similarity 92.4%; Pred. No. 1.2e-231;
Matches 502; Conservative 0; Mismatches 3; Indels 38; Gaps 1;

QY 1 MEGGPVCCDPAELVERVAALDVTHLEADGGPEPTRNGVDPFPPRARAASVIPGSTSR 60

Db 1 MEGGPVCCDPAELVERVAALDVTHLEADGGPEPTRNGVDPFPPRARAASVIPGSTSR 60

QY 61 LLPARSLARKLSLOERPAAGVLEAQAQYATGPAHSISPRAWRRPTTIESHVAISDAE 120

Db 61 LLPARSLARKLSLOERPAAGVLEAQAQYATGPAHSISPRAWRRPTTIESHVAISDAE 120

QY 121 DCVQLNOYKLOSEIGKAGVWRLAYNESDRHYAMKVLKLLKQYGFPPRPPRGSQ 180

Db 121 DCVQLNOYKLOSEIGKAGVWRLAYNESDRHYAMKVLKLLKQYGFPPRPPRGSQ 180

QY 181 AAGGPAKQLPLERVYQETAILKLDHNVVVKLIEVLDDPAEDNLYL----- 228

Db 181 AAGGPAKQLPLERVYQETAILKLDHNVVVKLIEVLDDPAEDNLYLALQNQAQNIQLD 240

QY 229 -----VFOLLRGPVMEVPCDKPFSEEQARLYLRDVIILG 262

Db 241 STNIAKSHSLSPSEQDSGTWAARSVFDLLRGPVMEVPCDKPFSEEQARLYLRDVIILG 300

QY 263 LEYLHCQIKVHRDIKPSNLLLGDDGHVKIADFGVSNQFEGNDAQLSSTAGTAPFMAPEAI 322

Db 301 LEYLHCQIKVHRDIKPSNLLLGDDGHVKIADFGVSNQFEGNDAQLSSTAGTAPFMAPEAI 360

QY 323 SDSGQSFSGKALDVWATGVTLYCFVYKCKPFIIDFIALHRKIKNEPVPVPEPPEISEEL 382

Db 361 SDSGQSFSGKALDVWATGVTLYCFVYKCKPFIIDFIALHRKIKNEPVPVPEPPEISEEL 420

QY 383 KDILKMLDKNPETRIGVDPDKLHPWTKNGEPLPSEEHCSVVEVTEGEVKNVRLIP 442

Db 421 KDILKMLDKNPETRIGVDPDKLHPWTKNGEPLPSEEHCSVVEVTEGEVKNVRLIP 480

QY 443 SWTTVILVKSMLRKRSFGNPFEPQARREERMSAPGNLLVKEGFGGKSPELPGVQEDE 502

Db 481 SWTTVILVKSMLRKRSFGNPFEPQARREERMSAPGNLLVKEGFGGKSPELPGVQEDE 540

QY 503 AAS 505

Db 541 AAS 543

RESULT 11

ABU65067

ID AEU65067 standard; protein; 503 AA.

AC ABU65067;

XX 20-MAY-2003 (first entry)

DT Human NOV15a protein.

DE NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;

KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;

KW human.

XX OS Homo sapiens.

XX PN WC200272757-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US006908.

XX PF

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR WPI; 2002-723332/78.
DR N-PSDB; ABX97034.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
XX asthma.
XX
XX Claim 1; Page 139; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
CC cytosolic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
XX Sequence 503 AA;
XX
Query Match 97.2%; Score 2560; DB 5; Length 503;
Best Local Similarity 98.2%; Pred. No. 2.2e-229;
Matches 496; Conservative 2; Mismatches 5; Indels 2; Gaps 2;
Qy 1 MEGGPAVCCDPRAEELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVPIGSTSR 60
Db 1 MEGGPAVCCDPRAEELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVPIGSTSR 60
Qy 61 LLPAPSLARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Db 61 LLPAPSLARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Qy 121 DCVQLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
Db 121 DCVQLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
Qy 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Db 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Qy 241 VPCDKPSEEARLYLRDVLILGLYHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
Db 240 VPCDKPSEEARLYLRDVLILGLYHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 299
Qy 301 EGNDALSSSTAGTAFMAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCPFFIDDFILA 360
Db 300 EGNDALSSSTAGTAFMAPEAISDSGQSFSGK-LDVWATGVTLYCFVYGKCPFFIDDFILA 358
Qy 361 LHRKIKNEPVVPPEPEISEELKOLILKMLDKNPETRIGVPDIKLPWVTXNGEPLPSE 420
Db 359 LHRKIKNEPVVPPEPEISEELKOLILKMLDKNPETRIGVPDIKLPWVTXNGEPLPSE 418
Qy 421 EEHCSVVEVTGEVKNVRLIPSTTTVILKSMLEKRSFGNPFEPQARREERSAPGNL 480
Db 419 EEHCSVVEVTGEVKNVRLIPSTTTVILKSMLEKRSFGNPFEPQARREERSAPGNL 478
Qy 481 LVKGFGEKGSPELPGVQDEAAS 505
Db 479 LVKGFGEKGSPELPGVQDEAAS 503
RESULT 12
ID AB883055
XX AB883055 standard; protein; 505 AA.
AC AB883055;
XX
XX 27-AUG-2002 (first entry)
XX Rat amino acid sequence related to the kinase protein of the invention.
XX

KW Rat; kinase protein; calcium/calmodulin-dependant protein kinase; eye;
KW retinoblastoma; brain; kinase modulator.
XX
XX Rattus norvegicus.
XX
XX WO2000224920-A2.
XX
XX 28-MAR-2002.
XX
XX 19-SEP-2001; 2001WO-US029161.
XX
XX 19-SEP-2000; 2000US-0233433P.
XX
XX 13-NOV-2000; 2000US-0247031P.
XX
XX 06-DEC-2000; 2000US-00729995.
XX
XX (PEKE) PE CORP NY.
XX
XX Beasley EM, Wei M, Bonazzi VR, Sanders R, Di Francesco V;
XX
XX WPI; 2002-404955/43.
XX
XX Novel peptide designated as human kinase useful as target for diagnosing
PT a disease or predisposition to the disease mediated by the peptide.
XX
XX Disclosure; Page 88-89; 89pp; English.
XX
XX The invention relates to an isolated peptide designated human kinase
CC (HK), that has homology to members of the calcium/calmodulin-dependent
CC protein kinase kinase subfamily. The mechanism of action of the protein
CC of the invention is that of a kinase modulator. The human kinase of the
CC invention is useful for creating a pharmaceutical composition for
CC treating a disease or condition mediated by the human kinase. HK is also
CC useful to provide a target for diagnosing a disease or predisposition to
CC disease mediated by HK, and is also useful in pharmacogenic analysis. HK
CC is useful for treating a disorder characterised by absence of
CC inappropriate or unwanted expression of HK, also as an immunogen to raise
CC antibodies by administering HK to a mammalian organism e.g. rat, rabbit
CC or mouse. Nucleic acids of the invention are useful as hybridisation
CC probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells
CC expressing a native HK are useful for assaying compounds that stimulate
CC or inhibit HK function. Nucleic acids of the invention are also useful
CC for producing transgenic animals. Experimental data indicates that kinase
CC proteins of the present invention are expressed in humans in the eye
CC (retinoblastomas) and brain. The current sequence represents a rat amino
CC acid sequence related to to kinase protein of the invention, appearing as
CC SEQ ID 4 in the sequence listing, but is not referred to in the
XX specification
XX Sequence 505 AA;
XX
Query Match 93.7%; Score 2468; DB 5; Length 505;
Best Local Similarity 93.5%; Pred. No. 8.5e-221;
Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Qy 1 MEGGPAVCCDPRAEELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVPIGSTSR 60
Db 1 MEGGPAVCCDPRAEELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVPIGSTSR 60
Qy 61 LLPAPSLARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Db 61 PTPVRLPSLSARKLSLOERPAGSCLEAQVGFSTGPASHMSPRAWRPTIESHHVAISDE 120
Qy 121 DCVQLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
Db 121 DCVQLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
Qy 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Db 181 APOGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Qy 241 VPCDKPSEEARLYLRDVLILGLYHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
Db 241 VPCDKPPEEARLYLRDVLILGLYHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300

QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPAFMAPEAISDTQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
QY 361 LHRKIKNEPVVFPPEEPEISEELKDILKMLDKNPETRIGVDPDIKLPWVTKNGEELPSE 420
DB 361 LHRKIKNEAVVFPPEEVESEELKDILKMLDKNPETRIGVSDIKLPWVTKNGEELPSE 420
QY 421 EHCSSVVEVTEGEVKNVRLIPSWTIVLYKSMRLKRSFGNPFEPQARREERSMSAPGNL 480
DB 421 EHCSSVVEVTEGEVKNVRLIPSWTIVLYKSMRLKRSFGNPFEPQARREERSMSAPGNL 480
QY 481 LVKEGFGGKGKSPELPGVQDEAAS 505
DB 481 LVKEGFGGKGKSPELPGVQDEAAS 505

RESULT 13

ID ADE61937
XX ADE61937 standard; protein; 505 AA.

AC ADE61937;
XX

DT 29-JAN-2004 (first entry)
XX

DE Rat Protein AAB46910, SEQ ID NO 7866.
XX

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX

OS Rattus norvegicus.
XX

PN WO2003016475-A2.
XX

PD 27-FEB-2003.
XX

XX 14-AUG-2002; 2002WO-US025765.
XX

XX 14-AUG-2001; 2001US-0312147P.
PR

PR 01-NOV-2001; 2001US-0346382P.
PR

PR 26-NOV-2001; 2001US-033347P.
XX

XX (GEO) GEN HOSPITAL CORP.
PA

PA (FARB) BAYER AG.
PA

XX Woolf C, D'urso D, Befort K, Costigan M;
PI

XX WPI; 2003-268312/26.
DR

XX GENBANK; AAB46910.
DR

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT

XX Claim 1; Page; 1017pp; English.
PS

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 505 AA;

Query Match 93.7%; Score 2458; DB 7; Length 505;
Best Local Similarity 93.5%; Pred. No. 8.5e-221;
Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY 1 MEGGPAVCCODPRAELVERVAADIVTHLEEAADGPEPTRNGVDPPPRARAASVIPGSTSR 60
DB 1 MERSPAVCCODPRAELVERVAALISVAHLEAEAGEPSPAGNVDPPPRARAASVIPGSASR 60

QY 61 LJPAPSLSARKLSLQERPAQSYLEAQAGFYATGPASHISPRAWRRPTTIESHHVAISDAE 120
DB 61 PTPVRPSLSARKPSLQERPAQSCLEAQGVFYSTGPASHMSPRAWRRPTTIESHHVAISDTE 120

QY 121 DCVQLNQYKLOSEIGKAGYGVRLAYNESEDRHYAMKVLKLLKQYGFPPRPPRGSQ 180
DB 121 DCVQLNQYKLOSEIGKAGYGVRLAYNEREDRHYAMKVLKLLKQYGFPPRPPRGSQ 180

QY 181 AAGGGPAKOLLPLERVYQEIATILKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
DB 181 APQGGPAKOLLPLERVYQEIATILKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240

QY 241 VPCDKPFSEEQARLYLRDVLGLVYHCOKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
DB 241 VPCDKPFSEEQARLYLRDILGLVYHCOKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300

QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPAFMAPEAISDTQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360

QY 361 LHRKIKNEPVVFPPEEPEISEELKDILKMLDKNPETRIGVDPDIKLPWVTKNGEELPSE 420
DB 361 LHRKIKNEAVVFPPEEPEVSEELKOLILKMLDKNPETRIGVSDIKLPWVTKNGEELPSE 420

QY 421 EHCSSVVEVTEGEVKNVRLIPSWTIVLYKSMRLKRSFGNPFEPQARREERSMSAPGNL 480
DB 421 EHCSSVVEVTEGEVKNVRLIPSWTIVLYKSMRLKRSFGNPFEPQARREERSMSAPGNL 480

QY 481 LVKEGFGGKGKSPELPGVQDEAAS 505
DB 481 LVKEGFGGKGKSPELPGVQDEAAS 505

RESULT 14

ADE61941

ID ADE61941 standard; protein; 505 AA.

XX ADE61941;
XX

XX 29-JAN-2004 (first entry)
DT

XX Rat Protein AAB46910, SEQ ID NO 7870.
DE

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX

OS Rattus norvegicus.
XX

PN WO2003016475-A2.
XX

XX 27-FEB-2003.
XX

PF	14-AUG-2002; 2002WO-US025765.	
XX		
PR	14-AUG-2001; 2001US-0312147P.	
PR	01-NOV-2001; 2001US-0346382P.	
PR	26-NOV-2001; 2001US-0333347P.	
XX		
XX	(GEO) GEN HOSPITAL CORP.	
PA	(FARB) BAYER AG.	
XX		
PI	WoOLF C, D'urso D, Befort K, Costigan M;	
XX		
DR	WPI; 2003-268312/26.	
DR	GENBANK; AAB46910.	
XX		
PT	New composition comprising two or more isolated polypeptides, useful for	
XX	preparing a medicament for treating pain in an animal.	
XX		
PS	Claim 1; Page; 1017pp; English.	
XX		
CC	The invention discloses a composition comprising two or more isolated rat	
CC	or human polynucleotides or a polynucleotide which represents a fragment,	
CC	derivative or allelic variation of the nucleic acid sequence. Also	
CC	claimed are a vector comprising the novel polynucleotide, a host cell	
CC	comprising the vector, a method for identifying a nucleotide sequence	
CC	which is differentially regulated in an animal subjected to pain and a	
CC	kit to perform the method, an array, a method for identifying an agent	
CC	that increases or decreases the expression of the polynucleotide sequence	
CC	that is differentially expressed in neuronal tissue of a first animal	
CC	subjected to pain, a method for identifying a compound which regulates	
CC	the expression of a polynucleotide sequence which is differentially	
CC	expressed in an animal subjected to pain, a method for identifying a	
CC	compound that regulates the activity of one or more of the	
CC	polynucleotides, a method for producing a pharmaceutical composition, a	
CC	method for identifying a compound or small molecule that regulates the	
CC	activity in an animal of one or more of the polypeptides given in the	
CC	specification, a method for identifying a compound useful in treating	
CC	pain and a pharmaceutical composition comprising the one or more	
CC	polypeptides or their antibodies. The polynucleotide or the compound that	
CC	modulates its activity is useful for preparing a medicament for treating	
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction	
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene	
CC	therapy). The sequence presented is a rat protein (shown in Table 2 of	
CC	the specification) which is differentially expressed during pain. Note:	
CC	The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic form directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 505 AA;	
	Query Match 93.7%; Score 2468; DB 7; Length 505;	
	Best Local Similarity 93.5%; Pred. NO. 8.5e-221;	
	Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;	
QY	1 MEGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTENGVDPPPRAAASVIFGSTR 60	
DB	1 MERSPAVCCQDPRAELVERVAISVAHLEEEAGEPEPASNGVDPPPRAAASVIFGSR 60	
QY	61 LLPARPSSLARKLSLOERPAGSYLEAQAQPYATGPAISHISPAWRPRTTIESHHVAISDAE 120	
DB	61 PTPVRPSSLARKFSLOERPAGSCLEAQGPYSTGPASHMSPAWRPRTTIESHHVAISDTE 120	
QY	121 DCVQLNQYKLOSEICKGAYGVVRLAYNSESDBHYAMKVLKKKLLIKOYGFRRPPRGSQ 180	
DB	121 DCVQLNQYKLOSEICKGAYGVVRLAYNERDRHYAMKVLKKKLLIKOYGFRRPPRGSQ 180	
QY	181 AAQGPAPKQLPLERVQEIATLKKLDHVVNVVKLIEVLDDPAEDNMLYVFDLLRKGPVME 240	
DB	181 APQGPAPKQLPLERVQEIATLKKLDHVVNVVKLIEVLDDPAEDNMLYVFDLLRKGPVME 240	
QY	241 VPCDKPFSEQARLYLRDVLGLLEYLHCQIKVHRDIKPSNLLIGDDGHVKIADFGVSNQF 300	
DB	241 VPCDKPFPEQARLYLRDIIILGLEYLHCQIKVHRDIKPSNLLIGDDGHVKIADFGVSNQF 300	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 05:38:15 ; Search time 5991 Seconds

(without alignments)

7551.528 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggaggggggtccagctgt.....aggaagcagggtgcattcc 1515

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

RST:*

1: em_estba:*

2: em_estum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1508.6	99.6	1518	29	AY416154 Homo sapi
2	1222.6	80.7	1518	29	AY416156 Mus muscu
3	1053.8	70.0	1445	29	AY416155 Pan trogl
4	940	62.0	1201	9	AL539375 AL539375

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	892.4	58.9	1034	13	EX401350
6	889.6	58.7	1201	13	EX395129
7	779.4	51.4	1096	12	BM805189
8	777.2	51.3	939	13	BU552890
9	742.4	49.0	880	12	EG705317
10	732	48.3	910	13	BQ932168
11	723	47.7	1066	13	BQ277875
12	722	47.7	972	13	BQ277875
13	692.6	45.7	893	12	BQ396486
14	680.6	44.9	896	13	EX328136
15	668.6	44.1	784	13	BQ571613
16	658.4	43.5	685	12	EG701043
17	646.6	42.7	786	12	B1819446
18	639.8	42.2	2508	11	AK031399
19	633.2	41.8	2854	11	AK032070
20	589.8	38.9	3139	11	AK044660
21	570.6	37.7	727	12	BM950087
22	556	36.7	1127	13	BU902211
23	555.8	36.7	749	12	B1756153
24	529.2	34.9	648	12	BQ294573
25	502.6	33.2	589	13	BU671280
26	468	30.9	636	12	B1546778
27	462.8	30.5	515	12	BM087446
28	439	29.0	900	12	B125425
29	429.8	28.4	841	13	BU115704
30	417	27.5	496	14	CB712914
31	411.4	27.2	672	13	BY735290
32	398.6	26.3	1201	9	AL582916
33	397.6	26.2	414	12	BM150448
34	395.4	26.1	915	10	BE562166
35	389.2	25.7	857	13	BU420206
36	381.4	25.2	894	13	BU367227
37	373.4	24.6	438	13	EX488903
38	371	24.5	503	12	BM951274
39	370.6	24.5	727	12	BM952005
40	368.8	24.3	700	10	BE887687
41	366.6	24.2	750	12	BM718658
42	365.2	24.1	1120	14	CD498154
43	362	23.9	905	13	EX350723
44	361	23.8	703	28	AQ629789
45	351.6	23.2	881	13	EX328334

ALIGNMENTS

RESULT 1	AY416154	Homo sapiens CAMK1 gene, partial sequence.	1518 bp	DNA linear	GSS 17-DEC-2003
LOCUS	AY416154	Genomic survey sequence.			
DEFINITION	AY416154.1	GI:39772114			
ACCESSION	AY416154				
VERSION	AY416154.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1518)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source Location/Qualifiers

1..1518

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>1518

/gene="CAMKK1"

/locus_tag="HCM5796"

ORIGIN

Query Match 99.6%; Score 1508.6; DB 29; Length 1518;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1511; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGGGGGTCCAGCTGTCTGCTCCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60

DB 1 ATGAGGGGGTCCAGCTGTCTGCTCCAGATCCTCGGCAGAGCTGGTAGAACGGGTG 60

QY 61 GCAGCCATCGATGTACTCTCTGGAGGAGCGAGATGCTGGCCCGCAGAGCCCTACTAGAAAC 120

DB 61 GCAGCCATCGATGTACTCTCTGGAGGAGCGAGATGCTGGCCCGCAGAGCCCTACTAGAAAC 120

QY 121 GTGTGTGAACCCCAACCAACGGGCGAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180

DB 121 GTGTGTGAACCCCAACCAACGGGCGAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180

QY 181 CTCTCTCCAGCCGGCTAGCCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 240

DB 181 CTCTCTCCAGCCGGCTAGCCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 240

QY 241 GGAAGCTATCTGAGGCGCAGGTGGCCCTTATGCCACGGGGCTGCGAGCCATCTCC 300

DB 241 GGAAGCTATCTGAGGCGCAGGTGGCCCTTATGCCACGGGGCTGCGAGCCATCTCC 300

QY 301 CCCGGGCTTGGGAGGCGCCAGTCCAGTCCCAACAGTGGCCATCTCAGATGCGAG 360

DB 301 CCCGGGCTTGGGAGGCGCCAGTCCAGTCCCAACAGTGGCCATCTCAGATGCGAG 360

QY 361 GACTGCGTGCAGTGAACAGTCAAGCTCAGAGTGCAGATGTCAGAGGTGCTACGGT 420

DB 361 GACTGCGTGCAGTGAACAGTCAAGCTCAGAGTGCAGATGTCAGAGGTGCTACGGT 420

QY 421 GTGGTGAAGCTGCGCTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480

DB 421 GTGGTGAAGCTGCGCTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480

QY 481 AAAAGAGTACTGAAGAGTATGGCTTTCCAGTGCCTCCCGAGAGGTCCAG 540

DB 481 AAAAGAGTACTGAAGAGTATGGCTTTCCAGTGCCTCCCGAGAGGTCCAG 540

QY 541 GCTGCCAGGAGGACCCAGCAGCAGTGTGCTCCCTGGAGCGGGGTACACGAGATT 600

DB 541 GCTGCCAGGAGGACCCAGCAGCAGTGTGCTCCCTGGAGCGGGGTACACGAGATT 600

QY 601 GCATCTGAAGAGTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTGGATGAC 660

DB 601 GCATCTGAAGAGTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTGGATGAC 660

QY 661 CCAGCTGAGCAACCTCTATTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 720

DB 661 CCAGCTGAGCAACCTCTATTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 720

QY 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGTGCCTCTACCTCGGGGAGCTCATC 780

DB 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGTGCCTCTACCTCGGGGAGCTCATC 780

QY 781 CTGGGCTCGAGTACTTGGCATCTGCCAGAGATGCTCCACAGGAGCATCAAGCCATCAAC 840

DB 781 CTGGGCTCGAGTACTTGGCATCTGCCAGAGATGCTCCACAGGAGCATCAAGCCATCAAC 840

QY 841 CTGCTCTCGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 900

DB 841 CTGCTCTCGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 900

QY 901 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGAAACCCAGCATTCATGGCCCCCGAG 960

DB 901 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGAAACCCAGCATTCATGGCCCCCGAG 960

QY 961 GCCATTTCTGATTTCCGGCCAGAGCTTCAGTGGGAAGSCCTTGGATGTATGGCCACTGGC 1020

DB 961 GCCATTTCTGATTTCCGGCCAGAGCTTCAGTGGGAAGSCCTTGGATGTATGGCCACTGGC 1020

QY 1021 GTACAGTTGTACTGCTTTGCTATGGGAAGTGCCTCATCGACAGATTTTCATCCTGGCC 1080

DB 1021 GTACAGTTGTACTGCTTTGCTATGGGAAGTGCCTCATCGACAGATTTTCATCCTGGCC 1080

QY 1081 CTCCACAGGAAGATCAAGATGAGCCGTGCTTTCTCAGGAGCCAGAAATCAGCGAG 1140

DB 1081 CTCCACAGGAAGATCAAGATGAGCCGTGCTTTCTCAGGAGCCAGAAATCAGCGAG 1140

QY 1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAAATCCCGAGACGAGATTTGGGGTG 1200

DB 1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAAATCCCGAGACGAGATTTGGGGTG 1200

QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCCCTTCGGAG 1260

DB 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCCCTTCGGAG 1260

QY 1261 GAGGAGCAGTGCAGCGTGGTGGAGTGCACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

DB 1261 GAGGAGCAGTGCAGCGTGGTGGAGTGCACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

QY 1321 ATCCGAGCTGGACCAAGTGTATCTGTTGAAGTCCATGTCTGAGGAAGCTTCCTTTGGG 1380

DB 1321 ATCCGAGCTGGACCAAGTGTATCTGTTGAAGTCCATGTCTGAGGAAGCTTCCTTTGGG 1380

QY 1381 AACCCGTTTACGCCCCAGCAGCGGAGGAGGATCCATGCTCTCCAGGAAACCTA 1440

DB 1381 AACCCGTTTACGCCCCAGCAGCGGAGGAGGATCCATGCTCTCCAGGAAACCTA 1440

QY 1441 CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGAA 1500

DB 1441 CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGAA 1500

QY 1501 GAGCAGGCTGCATCC 1515

DB 1501 GAGCAGGCTGCATCC 1515

RESULT 2

AY416156 1518 bp DNA linear GSS 17-DEC-2003

LOCUS Mus musculus CAMKK1 gene, VIRTUAL TRANSCRIPT, partial sequence.

DEFINITION genomic survey sequence.

ACCESSION AY416156

VERSION AY416156.1 GI:39772116

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1518)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ClarK,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1518)

AUTHORS ClarK,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,

JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 1445)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarialw,A., Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..1445 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598"
gene	<1..>1445 /gene="CAVKK1" /locus_tag="HCMS796"
ORIGIN	
Query Match	70.0%; Score 1059.8; DB 29; Length 1445;
Best Local Similarity	73.6%; Pred. No. 1.1e-223;
Matches 1064; Conservative	0; Mismatches 381; Indels 0; Gaps 0;
Qy	1 ATGAGGGGGTCCAGTGTCTGTGTCAGAGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
Db	1 ATGAGGGGGTCCAGTGTCTGTGTCANNANCTCGGCANNGCTGGTAGAACGGGTG 60
Qy	61 GCAGCCATCATGTGACTCTCTGGAGAGCGAGATGGTGGCCAGAGCTCTAGAAAC 120
Db	61 GCAGCCATCATGTGACTCTCTGGAGAGCGAGATGGTGGCCAGAGCTCTAGAAAC 120
Qy	121 GGTGTGACCCCCACCAACGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db	121 GGTGTGACCCCTCCACCAACGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Qy	181 CTGCTCCAGCGCGCTAGCTCTCAGCCAGGAAGCTTTCCCTPACAGAGCGGCCAGCA 240
Db	181 CTGCTCCAGCGCGCTAGNNTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Qy	241 GGAAGCTATCTGGAGGCGCAGGCTGGCCCTATGCCACGGGCTGCCAGCCATCTCC 300
Db	241 GGAAGCTATCTGGAGGCGCAGGCTGGCCCTATGCCACGGGCTGCCAGCCANNNN 300
Qy	301 CCCGGGCTCTGGGAGGCGCCACCATCGAGTCCACACCGTGGCCATCTCAGATGCAAG 360
Db	301 CCCGGGCTCTGNNAGGCGCCACCANNNAGTCCACACCGTGGCCANNNCAGATGCAAG 360
Qy	361 GACTGGTGCAGCTGAACCAAGTACAAAGCTGCAGAGTGAAGTGGCAAGGGTGCCTACGGT 420
Db	361 GACTGGTGCAGCTGAACCAAGTACAAAGCTGCAGAGTGAAGTGGCAAGGGTGCCTACGGT 420
Qy	421 GTGGTGAAGTGGCTTACAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTCC 480
Db	421 GTGGTGAAGTGGCTTACAAGAAAGTGAAGACAGACACTATNNNTGAAAGTCTTTCC 480
Qy	481 AAAAAAGTACTGAAGACGATATGGCTTTCCAGTCCGCTCCCGCAGAGGGTCCAG 540
Db	481 AAAAAAGTACTGAAGACGATATGGCTTTCCANNNNNNNNNNNNNNNNNNNNNNNN 540
Qy	541 GCTGCCAGGAGGACAGCAAGCAGCTGCTGCCCTGGAGCGGGTGTACCGAGATT 600
Db	541 NNN 600
Qy	601 GCCATCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGATCCAGAGTCTCGATGAC 660
Db	601 NNN 660
Qy	661 CCAGCTGAGGACAACTCTATTGTTGGTGTGACTCTGAGAAAGGGGCCCGCTCATGGAA 720
Db	661 NNN 720
Qy	721 GTGCCCTGTGACAAAGCCCTTTCGGAGGAGCAAGCTCGCCTTACCTGCGGAGCGTCATC 780
Db	721 NNN 780
Qy	781 CTGGCCCTCAGTACTTGGACTCCAGAAAGATCGTCCACAGGAGCATCAAGCCATCCAAC 840
Db	781 NNN 840
Qy	841 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 900
Db	841 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 900
Qy	901 GAGGGAAACGACGCTCAGCTGTCCAGCACCGCGGAAACCCGAGATTTCATGGCCCCGAG 960
Db	901 GAGGGAAACGACGCTCAGCTGTCCAGCACCGCGGAAACCCGAGATTTCATGGCCCCGAG 960
Qy	961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCCCTTGGATGATGGGCCACCGG 1020
Db	961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCCCTTGGATGATGGGCCACCGG 1020
Qy	1021 GTCAGCTGTACTGCTTGTATGGGAAGTGCCCATTCATCGACGATTTTCATCCTGGCC 1080
Db	1021 GTCAGCTGTACTGCTTGTATGGGAAGTNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
Qy	1081 CTCACAGGAAGATCAAGATGAGCCCGTGGTCTTCAGGAGCGAGAAATCAGCGAG 1140
Db	1081 NNN 1140
Qy	1141 GAGCTCAAGCACCTGATCTCTGAAGATGTTAGCAAGAAATCCCGAGAGAGAAATGGGGTG 1200
Db	1141 GAGCTCAAGCATCTGATCTCTGAAGATGTTAGCAAGAAATCCCGAGAGAGAAATGGGGTG 1200
Qy	1201 CCAGACATCAAGTTGGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1260
Db	1201 CCAGACATCAAGTTGGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1260
Qy	1261 GAGGAGACTGACAGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db	1261 GAGGAGACTGACAGCTGGTGGAGGTGACAGAGAGAGAGGTTAAGAACTCAGTCAGGCTC 1320
Qy	1321 ATCCAGCTGGACCAAGCTGCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1380
Db	1321 ATCCANN 1380
Qy	1381 AACCCGTTGAGCCCGACAGGACAGGAAAGCGATCCATGTCTGCTCCAGGAAACCTA 1440
Db	1381 AACCCGTTGAGCCCGACAGGACAGGAAAGCGATCCATGTCTGCTCCAGGAAACCTA 1440
Qy	1441 CTGGT 1445
Db	1441 CTGGT 1445
RESULT 4	
AL539375	
LOCUS	AL539375 Homo sapiens FETAL BRAIN Homo sapiens cdna clone
DEFINITION	CS0DF033Y017 5-PRIME, mRNA sequence.
ACCESSION	AL539375
VERSION	AL539375.2 GI:31263943
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12868532. Contact: Genoscope Genoscope - Centre National de Sequencage

[illegible]

Qy	1025 CGTGTACTGCTTT 1038 1078 CGTGTACTGHTTK 1091
RESULT 5	
EX401350	
LOCUS	EX401350 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens linear EST 13-MAY-2003
DEFINITION	cDNA clone CS0DK012YG22 5-PRIME, mRNA sequence.
ACCESSION	BX401350
VERSION	BX401350
KEYWORDS	EST.
SOURCE	BX401350.1 GI:30630361
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1. (Bases 1 to 1034)
COMMENT	Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8786.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DK012BD11QP1&cluster=8786.r . Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK012BD11QP1. Location/Qualifiers 1. .1034 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DK012YG22" /cell_line="HELA CELLS COT 25-NORMALIZED" /clone_lib="HELA" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES	
source	
ORIGIN	

Query Match 58.9%; Score 892.4; DB 13; Length 1034;
Best Local Similarity 99.2%; Pred. No. 1,1e-186;
Matches 904; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATGGAGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGTGTAGAACGGGTG 60
Db |||||
QY 61 GCAGCCATCGATGTGACTCACTTGCAGAGGAGGAGATGTGGCCACAGAGCTACTAGAAC 120
Db |||||
QY 157 GCAGCCATCGATGTGACTCACTTGCAGAGGAGGAGATGTGGCCACAGAGCTACTAGAAC 216
Db |||||
QY 121 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db |||||
QY 217 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 276
Db |||||
QY 181 CTGCTCCAGCCCGGCTTAGCTCTCAGCAGGAAGCTTTCCCTACAGAGCGGCAGCA 240
Db |||||
QY 277 CTGCTCCAGCCCGGCTTAGCTCTCAGCAGGAAGCTTTCCCTACAGAGCGGCAGCA 336
Db |||||
QY 241 GGAAGCTATCTGGAGGCGCAGCTGGGCTTTATGCCAGGGGCTGCCAGCCACATCTCC 300
Db |||||
QY 337 GGAAGCTATCTGGAGGCGCAGCTGGGCTTTATGCCAGGGGCTGCCAGCCACATCTCC 396
Db |||||
QY 301 CCCCGGGCTCGCGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db |||||
QY 397 CCCCGGGCTCGCGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 456
Db |||||
QY 361 GACTGCTGTCAGCTGAACACAGTACAGTGCAGAGTGAATGTGGAAGGGTGCCTACGGT 420
Db |||||
QY 457 GACTGCTGTCAGCTGAACACAGTACAGTGCAGAGTGAATGTGGAAGGGTGCCTACGGT 516
Db |||||
QY 421 GTGTGTAGGCTGGCTTACAAAGTGAACAGACAGACACTGAATGAAGTCTTCTTCC 480
Db |||||
QY 517 GTGTGTAGGCTGGCTTACAAAGTGAACAGACAGACACTGAATGAAGTCTTCTTCC 576
Db |||||
QY 481 AAAAAGAGTTACTGAACAGTATGGCTTTCCAGTGGCCCTCCCGCAGAGGGTCCAG 540
Db |||||
QY 577 AAAAAGAGTTACTGAACAGTATGGCTTTCCAGTGGCCCTCCCGCAGAGGGTCCAG 636
Db |||||
QY 541 GCTGCCAGGAGGACAGCAGCTGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 600
Db |||||
QY 637 GCTGCCAGGAGGACAGCAGCTGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 696
Db |||||
QY 601 GCATCTCTGAAGAGCTGGACCACTGAATGTGTCAAACCTGATCGAGGCTCTGATGAC 660
Db |||||
QY 697 GCATCTCTGAAGAGCTGGACCACTGAATGTGTCAAACCTGATCGAGGCTCTGATGAC 756
Db |||||
QY 661 CCAGCTGAGGACAACTCTATTGCTTTTGAACCTCTCTGAGAAAGGGGCCCGTCAATGAA 720
Db |||||
QY 757 CCAGCTGAGGACAACTCTATTGCTTTTGAACCTCTCTGAGAAAGGGGCCCGTCAATGAA 816
Db |||||
QY 721 GTGCCCTGTGACAGCCCTTCTCGAGAGCAGCTGCTGCTTACCTCGGGAGCTGATC 780
Db |||||
QY 817 GTGCCCTGTGACAGCCCTTCTCGAGAGCAGCTGCTGCTTACCTCGGGAGCTGATC 876
Db |||||
QY 781 CTGGGCTCGAGTACTTGTCAATGTCAGAGATCTCCACAGGACATCAAGCCATCCAAC 840
Db |||||
QY 877 CTGGGCTCGAGTACTTGTCAATGTCAGAGATCTCCACAGGACATCAAGCCATCCAAC 936
Db |||||
QY 841 CTGCTCTT-GGGGATGATGGGCACTGAAGATCGCCGATTTGGGCTCGAGAACCACTT 899
Db |||||
QY 937 CTGCTCTTGGGGGATGATGGGCACTGAAGATCGCCGACTTTGGCGTCAACACCMAGTT 996
Db |||||
QY 900 TGAGGGGAACG 910
Db |||||
QY 997 TGAGGGGAACR 1007
Db |||||

RESULT 6
BX395129 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX395129 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION

CDNA clone CS0DD001YE15 5-PRIME, mRNA sequence.
BX395129
VERSION BX395129.1 GI:30628440
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8786.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD001AC08QPl&cluster=8786.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVivoGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD001AC08QPl.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD001YE15"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dr)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 58.7%; Score 889.6; DB 13; Length 1201;
Best Local Similarity 99.0%; Pred. No. 4.8e-186;
Matches 899; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAGAGGGGTTCAGCTGTCTGCTGCCAGATCTCGGCAGAGCTGTGGTAGAACGGGTG 60
Db |||||
QY 83 ATGAGAGGGGTTCAGCTGTCTGCTGCCAGTCTCGGCAGAGCTGTGGTAGAACGGGTG 142
Db |||||
QY 61 GCAGCCATCGATGTGACTCACTTTGGAGGAGCAGATGTGGCCAGAGCTTACTAGAAC 120
Db |||||
QY 143 GCAGCCATCGATGTGACTCACTTTGGAGGAGCAGATGTGGCCAGAGCTTACTAGAAC 202
Db |||||
QY 121 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGAGTACTTCAAGA 180
Db |||||
QY 203 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGAGTACTTCAAGA 262
Db |||||
QY 181 CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db |||||
QY 263 CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 322
Db |||||
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTTATGCCACGGGGCTGCCAGCCCATCTCC 300
Db |||||
QY 323 GGAAGCTATCTGAGGCGCAGGCTGGGCTTTATGCCACGGGGCTTCBAGCBACATCTCC 382
Db |||||
QY 301 CCCCGGGCTGGCGAGGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db |||||
QY 383 CCCCGGGCTGGCGAGGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 442
Db |||||
QY 361 GACTGCTGCAGCTGAACCAAGTCAAGCTCAGAGTGAATGTGGAGGGTGCCTACGGT 420
Db |||||
QY 443 GACTGCTGCAGCTGAACCAAGTCAAGCTCAGAGTGAATGTGGAGGGTGCCTACGGT 502
Db |||||
QY 421 GTGGTGAAGCTGCTTACACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTCC 480
Db |||||
QY 503 GTGGTGAAGCTGCTTACACGAAAGTGAAGACAGACTATGCAATGAAGTCTTTTCC 562
Db |||||

cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

QY	541	GTGTGCCAGGGAGGACCGACAGAGCTGTGTGCCCTGGAGCGGGTGTACGAGGATT	600
Db	623	GCTGCCAGGGAGGACCGACAGAGCTGTGTGCCCTGGAGCGGGTGTACGAGGATT	682
QY	601	GCATCTCTGAAGAAGCTGACACACGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC	660
Db	683	GCATCTCTGAAGAAGCTGACACACGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC	742
QY	661	CCAGCTGAGGACAAACCTCTATTATTGGTGTGTGACCTCTTGAGAAAGGGGCCCGTCAATGAA	720
Db	743	CCAGCTGAGGACAAACCTCTATTATTGGTGTGTGACCTCTTGAGAAAGGGGCCCGTCAATGAA	802
QY	721	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCCTTACCTCTCGGGAGCGTCAATC	780
Db	803	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCCTTACCTCTCGGGAGCGTCAATC	862
QY	781	CTGGGCTCTGAGTACTTGCACTGCCAGAGATCGTCCACAGGAGCATCAAGCCATCCAAC	840
Db	863	CTGGGCTCTGAGTACTTGCACTGCCAGAGATCGTCCACAGGAGCATCAAGCCATCCAAC	922
QY	841	CTGCTCTGGGGATGATGGGCACGTGAGATCGCGGACTTTGGCGTCAGCAACACGATTT	900
Db	923	CTGCTCTGGGGATGATGGGCACGTGAGATCGCGGACTTTGGCGTCA-SAACCAAGTTT	981
QY	901	GAGGGGAA 908	
Db	982	GAGGGGAA 989	
RESULT 7			
BM805189			
LOCUS			
DEFINITION			
AGENCOURT_6499735 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5728398			
5', mRNA sequence.			
BM805189			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1096)			
NIH-MGC http://mgi.nci.nih.gov/ .			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: Invitrogen			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
Cloning by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM12724 row: a column: 07			
High quality sequence stop: 633.			
Location/Qualifiers			
1. 1096			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5728398"			
/tissue_type="hippocampus"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_124"			
/note="Organ: brain; Vector: pCMV-sPORT6; Site 1: EcoRV			
(destroyed); Site 2: NotI; RNA source male hippocampus,			
age 27. Library is oligo-dT primed and directionally			
source			
FEATURES			
Query Match 51.4%; Score 779.4; DB 12; Length 1096;			
Best Local Similarity 93.1%; Pred. No. 1.2e-161;			
Matches 905; Conservative 0; Mismatches 52; Indels 15; Gaps 8;			
QY	1	ATGGAGGGGGTCCAGCTGTGTGTGCCAGGATCTCTGGGCGAGAGCTGGTAGAACGGGTG	60
Db	78	ATGGAGGGGGTCCAGCTGTGTGTGCCAGGATCTCTGGGCGAGAGCTGGTAGAACGGGTG	137
QY	61	GCAGCCATCGATGTGACTCACTTTGGAGGAGGACAGATGGTGGCCGAGAGCTTACTAGAAC	120
Db	138	GCAGCCATCGATGTGACTCACTTTGGAGGAGGACAGATGGTGGCCGAGAGCTTACTAGAAC	197
QY	121	GGTGTGGACCCGCCACACGAGGGCCAGAGCTGCTGTGTGATCCCTGCGAGTACTTCAAGA	180
Db	198	GGTGTGGACCCGCCACACGAGGGCCAGAGCTGCTGTGTGATCCCTGCGAGTACTTCAAGA	257
QY	181	CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA	317
Db	258	CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA	377
QY	241	GGAGCTATCTGGAGGCGCAGCTGGGCTTTATGCCAGGGGCTGCCAGCCATCTCC	300
Db	318	GGAGCTATCTGGAGGCGCAGCTGGGCTTTATGCCAGGGGCTGCCAGCCATCTCC	377
QY	301	CCCGGGCTGGCGGAGGCCACCATCGAGTCCCACACGCTGGCCATCTCAGATCGCAG	360
Db	378	CCCGGGCTGGCGGAGGCCACCATCGAGTCCCACACGCTGGCCATCTCAGATCGCAG	437
QY	361	GACTGCTGACGTGAACACAGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT	420
Db	438	GACTGCTGACGTGAACACAGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT	497
QY	421	GTGCTGAGGCTGGCCCTACAAACGAAAAGTGAAGACACACTATGCAATGAAAAGTCTCTTCC	480
Db	498	GTGCTGAGGCTGGCCCTACAAACGAAAAGTGAAGACACACTATGCAATGAAAAGTCTCTTCC	557
QY	481	AAAAAGAAATTACTGAAGCAGTATGGCTTTCACGTGCGCTTCCCCGAGAGGGTCCCGAG	540
Db	558	AAAAAGAAATTACTGAAGCAGTATGGCTTTCACGTGCGCTTCCCCGAGAGGGTCCCGAG	617
QY	541	GCTGCCAGGGAGGACCGCCAGCAGCTGCTG-CCCTGGAGCGGGTGTACAGGAGAT	599
Db	618	GCTGCCAGGGAGGACCGCCAGCAGCTGCTG-CCCTGGAGCGGGTGTACAGGAGAT	677
QY	600	TGCCATCTCTGAAGAGCTGGACCACTGAATGTGTGTAACCTGATCGAGGTCTTGATGA	659
Db	678	TGCCATCTCTGAAGAGCTGGACCACTGAATGTGTGTAACCTGATCGAGGTCTTGATGA	737
QY	660	CCAGCTGAGGACAACTCTATTGG-TGTTGACTCTCTGAGAAAGGG-CCCGTCAATG	717
Db	738	CCAGCTGAGGACCACTCTATTGGTGTGTTGACTCTCTGAGAAAGGGGCCCGTCAATG	797
QY	718	GAAAGTGCCTGTGACAA-GCCCTTCTCGGAGGAGCAAGCTCGCTCTACCTCGCGGACGT	776
Db	798	GAAAGTGCCTGTGACAAAGGCCCTTCTCGGAGGAGCAAGCTCGCTCTACCTCGCGGACGT	857
QY	777	CATCC-TGGGCTCGAGTACTTGCACTGCCAGAGATGTCACAGGAGACATCAAGCCAT	835
Db	858	CATCTCTGGGCCCTCGAGTACTTGCACTGCCAGAGATGTCACAGGAGACATTAAGCCAT	917
QY	836	CCAACCTGCTCTGGGG-----GATGATGGGCACTGAAGATCGCCGACTTTGGCGTCA	890
Db	918	TCCAACCTGGTTCTTGGGGGGATGATGGGGACCTGAAGATCGCCCAACTTTTGGGGGCA	977
QY	891	CAACC--AGTTTGAGGGGAACGACGCTAGCTGTCC---AGCACGGCGGAACCCACCA	945

/clone_id="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally

```
Db 978 CAACCCANTTTGAGGGGAACACACGCTCAACTGTTCCAGAACGGGGGGAACCCCAACA 1037
QY 946 TTCTATGGCCCC 957
Db 1038 TTTCTGGCCCC 1049

RESULT 8
LOCUS BU552890
DEFINITION AGENCOURT_10373297 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6576737 5', mRNA sequence.
ACCESSION BU552890
VERSION BU552890.1 GI:22903162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2775 row: 1 column: 17
High quality sequence stop: 658.
FEATURES
source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6576737"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/notes="Organ: Ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 51.3%; Score 777.2; DB 13; Length 939;
Best Local Similarity 98.7%; Pred. No. 3.3e-161;
Matches 815; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY 685 GTGTTTGACCTCTGTGAAAGGGGCGCTCATGAGAGTGCCTGTGACAGCCCTTCTCG 744
Db 38 GTGTTTGACCTCTGTGAAAGGGGCGCTCATGAGAGTGCCTGTGACAGCCCTTCTCG 97
QY 745 GAGGAGCAAGCTCGCTCTACTCTGGGGAGCTCATCTGGGGCTCGAGTACTTGCACCTGC 804
Db 98 GAGGAGCAAGCTCGCTCTACTCTGGGGAGCTCATCTGGGGCTCGAGTACTTGCACCTGC 157
QY 805 CAGAAGATCGTCCACAGGAGACATCAAGCCATCAACCTGCTCTGGGGATGATGGGCAC 864
Db 158 CAGAAGATCGTCCACAGGAGACATCAAGCCATCAACCTGCTCTGGGGATGATGGGCAC 217
QY 865 GTGAAGATCGCCGACTTTGGCGTTCAGCAACACAGTTTTCAGGGGGAACACGCTCAGCTGTCC 924
Db 218 GTGAAGATCGCCGACTTTGGCGTTCAGCAACACAGTTTTCAGGGGGAACACGCTCAGCTGTCC 277
```

```
QY 925 AGCAGCGGGGAACCCAGCATTTCATGGCCCCCGAGGCCATTTCGATTCCGGCCAGAGC 984
Db 278 AGCAGCGGGGAACCCAGCATTTCATGGCCCCCGAGGCCATTTCGATTCCGGCCAGAGC 337
QY 985 TTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGCGTCACGTTCTACTGCTTCTCTAT 1044
Db 338 TTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGCGTCACGTTCTACTGCTTCTCTAT 397
QY 1045 GGGAAAGTCCCATTCATCGACGATTTCATCTCGCCCTCCACAGGAAGATCAAGAATGAG 1104
Db 398 GGGAAAGTCCCGTTTCATCGACGATTTCATCTCGCCCTCCACAGGAAGATCAAGAATGAG 457
QY 1105 CCCGTGCTGTTTCTCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTCTGAAG 1164
Db 458 CCCGTGCTGTTTCTCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTCTGAAG 517
QY 1165 ATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTGCCAGACATCAAGTTCCACCCCTGG 1224
Db 518 ATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTGCCAGACATCAAGTTCCACCCCTGG 577
QY 1225 GTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGGAG 1284
Db 578 GTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGGAG 637
QY 1285 GTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCAGCTGACACCGTGTATC 1344
Db 638 GTGACAGAGGAGGAGGTTAAGAACTCAGTCAGGCTCATCCAGCTGACACCGTGTATC 697
QY 1345 CTGCTGAAGTCCATGCTGAGGAGCGTTCCTTTGGGAACCCGTTTGAGCCGCCAGGACGG 1404
Db 698 CTGCTGAAGTCCATGCTGAGGAGCGTTCCTTTGGGAACCCGTTTGAGCCGCCAGGACGG 757
QY 1405 AGGGAAGAGCGATCCATGCTCTCCAGGAAACCTA-CTGGTGAAGAGAGG-TTGGTG 1462
Db 758 AGGGAAGAGCGATCCATGCTCTCCAGGAAACCTA-CTGGTGAAGAGAGG-TTGGTG 817
QY 1463 AAGGGGCAAGAGCCAGAGCTCCCGGC-GTCCAGGAAGACGAGG 1507
Db 818 AAGGGGCAAGAGCCCAAGACTCCCGCGGCTCCAGGAAGACCAAG 863
```

RESULT 9

```
LOCUS BG705317
DEFINITION 602687825F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4820238 5',
mRNA sequence.
ACCESSION BG705317
VERSION BG705317.1 GI:13979532
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10725 row: a column: 07
High quality sequence stop: 769.
FEATURES
source
1..880
/organism="Homo sapiens"
/mol_type="mRNA"
```

/db_xref="taxon:9606"
/clone="IMAGE:4820238"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: Brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.0%; Score 742.4; DB 12; Length 880;
Best Local Similarity 95.7%; Pred. No. 1.6e-153;
Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 6;

QY 26 GCCAGATCCTCGGCAGAGCTGTAGAACGGGTGGCAGCCATCGATGACTCACTTGG 85
DB 4 GGCAGATCCTCGGCAGAGCTGTAGAACGGGTGGCAGCCATCGATGACTCACTTGG 63

QY 86 AGGAGCAGATGCTGCCAGAGCTTACTAGAAACGGTGTGACACCCGCCACACGCGCCA 145
DB 64 AGGAGCAGATGCTGCCAGAGCTTACTAGAAACGGTGTGACACCCGCCACACGCGCCA 123

QY 146 GAGTGCCTCTGTGATCCCTGCGAGTACTCAAGACTGCTCCAGCCCGGCTAGCCTT 205
DB 124 GAGTGCCTCTGTGATCCCTGCGAGTACTCAAGACTGCTCCAGCCCGGCTAGCCTT 183

QY 206 CAGCAGGAGCTTCCCTACAGGCGGCCAGCAGGAGCTATCTGGAGGCGCAGGCTG 265
DB 184 CAGCAGGAGCTTCCCTACAGGCGGCCAGCAGGAGCTATCTGGAGGCGCAGGCTG 243

QY 266 GGCCTTATGCCAGGGGCTGCCAGCACATCTCCCGCGGGCTGGCGAGGCCACCA 325
DB 244 GGCCTTATGCCAGGGGCTGCCAGCACATCTCCCGCGGGCTGGCGAGGCCACCA 303

QY 326 TCGAGTCCACACGTCGGCCATCTCAGATCGAGAGACTGCGTGCAGCTGAACAGTACA 385
DB 304 TCGAGTCCACACGTCGGCCATCTCAGATCGAGAGACTGCGTGCAGCTGAACAGTACA 363

QY 386 AGCTGCAGAGTGAAGTGGCAAGGGTCTACGGTGTGTGAGGCTGGCTACACGAAA 445
DB 364 AGCTGCAGAGTGAAGTGGCAAGGGTCTACGGTGTGTGAGGCTGGCTACACGAAA 423

QY 446 GTCAAGACAGACTATGCAATGAAAGTCCCTTTCCAAAAGAGTTACTGAAGCAGTATG 505
DB 424 GTCAAGACAGACTATGCAATGAAAGTCCCTTTCCAAAAGAGTTACTGAAGCAGTATG 483

QY 506 GCTTTCCAGTGCCTCTCCCGAGAGGGTCCAGAGTCCCGAGGAGCAGCAGCAGCAAGC 565
DB 484 GCTTTCCAGTGCCTCTCCCGAGAGGGTCCAGAGTCCCGAGGAGCAGCAGCAGCAAGC 543

QY 566 AGCTGTCGCCCTGGAGCGGCTTACAGAGATTGCCATCTCTGAGAGAGCTGGACCACG 625
DB 544 AGCTGTCGCCCTGGAGCGGCTTACAGAGATTG-CATCTGAGAGAGCTGGACCACG 602

QY 626 TGAATGTGTCAAACTGATCAGGTCTGTGATGACCCAGCTGAGGACAACTCTATTGG 685
DB 603 TGAATGTGTCAAACTGATCAGGT-CTGATGACCCAGCTGAGGACAACTCTATTGG 661

QY 686 TGTTTGACCTCTTGAGAAAGGGCCCGTCAAGAGTGCCTCTGTGACAAAGCCCTTCTGG 745
DB 662 TG-TTGACCTCTTGAGAAAGGGG-CGCTCATGAAAGTGCCCTGTGACAAAGCCCTTCTGG 719

QY 746 AGGAGCAGCTCGCTCTACTCGGGGACGTCATCTCTGGGCTCGAGTACTGCACTGCC 805
DB 720 AGGAGCAGCTCGCTCTACTCGGGGACGTCATCTCTGGGCTCGAGTACTTGGACTGCC 779

QY 806 AGAA-GATCGTCCACA-GGGACATCAAGCCATCCAACTGCTCTCTGGGGATGATGGCA 863
DB 780 AGAAGATCGTCCACAGGGGACCTCAGCGCATGCAACCTGCTCTGGGGGATGATGGCC 839

QY 864 CQTGAGATGCCGACTTTGGCGT 887
DB 840 GGTAAAGCATCGGCAAGTAGGCGGT 863

RESULT 10
BQ932168
LOCUS
DEFINITION
IMAGE:6196986 5', mRNA sequence.
BQ932168
ACCESSION
VERSION
BQ932168.1 GI:22347199
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
1 (bases 1 to 910)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13606 row: e column: 19
High quality sequence stop: 577.
Location/Qualifiers
1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6196986"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGTCCG-3' and 5'-GACTAGTCTTAGATCGGAGCGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 48.3%; Score 732; DB 13; Length 910;
Best Local Similarity 94.4%; Pred. No. 3.4e-151;
Matches 849; Conservative 0; Mismatches 31; Indels 19; Gaps 8;

QY 100 GGCCCCAGAGCTACTAGAAACGGTGTGACCCGCCACACGAGCTGCTCTGTG 159
DB 1 GGCCCCAGAGCTACTAGAAACGGTGTGACCCGCCACACGAGCTGCTCTGTG 60

QY 160 ATCCCTGGCAGTACTTCAAGACTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTT 219
DB 61 ATCCCTGGCAGTACTTCAAGACTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTT 120

QY 220 TCCTTACAGAGCGGCCAGAGGAGCTATCTGGAGGCGCAGGCTGGGCTTATGCACG 279


```
Db 121 TCCCTACAGGAGCGGCCAGCAGGAAGCTATCTGGAGGCGCAGGCTGGGCTTTATGCCAGG 180
Qy 280 GGGCTGCGCAGCCACATCTCCCGCGGCTGGCGGAGGCCACCATGAGTCCACACAC 339
Db 181 GGGCTGCGCAGCCACATCTCCCGCGGCTGGCGGAGGCCACCATGAGTCCACACAC 240
Qy 340 GTGGCCATCTCAGATGCAGAGACTGCGTGCAGCTGAACCACTACAAGTCAGAGTGAG 399
Db 241 GTGGCCATCTCAGATGCAGAGACTGCGTGCAGCTGAACCACTACAAGTCAGAGTGAG 300
Qy 400 ATTGCAAGGCTGCTCAGTGTGTGTGGCTGGCTTCAACGAAAGTGAAGACAGACAC 459
Db 301 ATTGCAAGGCTGCTCAGTGTGTGTGGCTGGCTTCAACGAAAGTGAAGACAGACAC 360
Qy 460 TATGCAAGGCTGCTCAGTGTGTGTGGCTGGCTTCAACGAAAGTGAAGACAGACAC 519
Db 361 TATGCAAGGCTGCTCAGTGTGTGTGGCTGGCTTCAACGAAAGTGAAGACAGACAC 420
Qy 520 CTTCCCGGAGAGGCTCCAGGCTGCCAGGAGGAGCAGCAGCAGCTGCTGCCCTG 579
Db 421 CTTCCCGGAGAGGCTCCAGGCTGCCAGGAGGAGCAGCAGCAGCTGCTGCCCTG 480
Qy 580 GAGCGGTGTACAGAGATTCGCAATCTGCAAGAGTGGACCAAGTGTGTGTGCTCAA 639
Db 481 GAGCGGTGTACAGAGATTCGCAATCTGCAAGAGTGGACCAAGTGTGTGTGCTCAA 540
Qy 640 CTGATCGAGGTCCTGGATGACCCAGCTGAGGACACCTCTATTTGGTGTGAGCTCTG 699
Db 541 CTGATCGAGGTCCTGGATGACCCAGCTGAGGACACCTCTATTTGGTGTGAGCTCTG 600
Qy 700 AGAAAGGCGGCGCTCATGGAAGTGCCTGTGCAAGCCCTTCTCGGAGGAGC-AAAGCTCG 758
Db 601 AGAAAGGCGGCGCTCATGGAAGTGCCTGTGCAAGCCCTTCTCGGAGGAGC-AAAGCTCG 660
Qy 759 CTTTACTCGGGAAGTCTCTGGGCTCGAGTCTGACACTGTCAGTGTGCCAGAGATCGTCCA 818
Db 661 CTTTACTCGGGAAGTCTCTGGGCTCGAGTCTGACACTGTCAGTGTGCCAGAGATCGTCCA 720
Qy 819 CA-GGGACATCAAGCCATCAACCTGCTCC--TGGGGGATGATGGGACGTCAGATCG 874
Db 721 CAGGGGACATCAAGCCATCAACCTGCTCC--TGGGGGATGATGGGACGTCAGATCG 780
Qy 875 CCGAC-----TTTGGCGTCAGCAACCCAGTTT---GAGGGGAAACGAGCTCAGC-TGTCC 924
Db 781 TCGCCCGACATTTGGCGTCAGCAACCCAGTTT---GAGGGGAAACGAGCTCAGC-TGTCC 840
Qy 925 AGCAGCGGGGAA--CCCGACATTCATGG--CCCGCAGGCGCATTTCTGATTCGGGCC 979
Db 841 AGCAACGCGGAAACCCCGACATTCATGGGCCCCCGAGGCGCATTTCTGAATTCGGGC 899
```

```
RESULT 11
B0277875
LOCUS B0277875 1066 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_7048118 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5804754
5', mRNA sequence.
ACCESSION B0277875
VERSION B0277875.1 GI:20488083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1066)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LJNL)
DNA Sequencing by: Agencourt Bioscience Corporation
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LJNL at: <http://image.llnl.gov>
Plate: LCM2042 row: n column: 19
High quality sequence stop: 626.
Location/Qualifiers
source
1..1066
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5804754"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

ORIGIN

```
Query Match 47.7%; Score 723; DB 13; Length 1066;
Best Local Similarity 97.3%; Pred. No. 3.6e-149;
Matches 779; Conservative 0; Mismatches 15; Indels 7; Gaps 4;
Qy 685 GTGTTTACCTCTCTGAGAAAGGGCGGCTCATGCAAGTGCCTGTGCAAGCCCTTCTCG 744
Db 36 GTGTTTACCTCTCTGAGAAAGGGCGGCTCATGCAAGTGCCTGTGCAAGCCCTTCTCG 95
Qy 745 GAGGAGCAGCTCGCTCTTACCTCGGAGCTCATCTCGGCTTCGAGTACTTGCACTGC 804
Db 96 GAGGAGCAGCTCGCTCTTACCTCGGAGCTCATCTCGGCTTCGAGTACTTGCACTGC 155
Qy 805 CAGAAGATCGTCCACAGGACATCAAGCCATCCAACTGCTTCTCGGAGTATGGGCAC 864
Db 156 CAGAAGATCGTCCACAGGACATCAAGCCATCCAACTGCTTCTCGGAGTATGGGCAC 215
Qy 865 GTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTTTCAGGGGAAACGAGCTCAGCTGTCC 924
Db 216 GTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTTTCAGGGGAAACGAGCTCAGCTGTCC 275
Qy 925 AGCAGCGGGGAAACCCAGCATTCATGCGCCCGGAGGCCATTTCTGATTCGGGCAAGC 984
Db 276 AGCAGCGGGGAAACCCAGCATTCATGCGCCCGGAGGCCATTTCTGATTCGGGCAAGC 335
Qy 985 TTCAGTGGGAAAGCCTTGGATGTATGGGCGACATGGCGTACGTTGCTGTTGCTAT 1044
Db 336 TTCAGTGGGAAAGCCTTGGATGTATGGGCGACATGGCGTACGTTGCTGTTGCTAT 395
Qy 1045 GGAAGTGGCCATTCATTCGACGATTTTCATCTGGCCCTCCACAGGAAGATCAAGATGAG 1104
Db 396 GGAAGTGGCCGTTTCATTCGACGATTTTCATCTGGCCCTCCACAGGAAGATCAAGATGAG 455
Qy 1105 CCGTGTGTGTTCTCGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTGAAG 1164
Db 456 CCGTGTGTGTTCTCGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTGAAG 515
Qy 1165 ATGTTAGCAAGAATCCGAGAGAGAGATTTGGGTCGACACATCAAGTTGACCCCTGG 1224
Db 516 ATGTTAGCAAGAATCCGAGAGAGAGATTTGGGTCGACACATCAAGTTGACCCCTGG 575
Qy 1225 GTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAGGAGGACATTCAGCGTGTGGAG 1284
Db 576 GTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAGGAGGACATTCAGCGTGTGGAG 635
Qy 1285 GTGACAGAGGGGAGGTTAAGACTCAGTCAGGCTCATCCGAGCTGACACCGGTGATC 1344
Db 636 GTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCGAGCTGGAACGAGGTGATC 695
Qy 1345 CTGGTGAAGTCCATGCTG-AGGAAGCGTTTCTTCTGGGAACCCCTTTGAGGCCAGGACG 1403
```

```
Db      696  CTGGGGAAGTCCATGCTGAAGGAAGCGTTTCCTTTGGGAACCCCTTTGAGCCCAAGCAG 755
      1404  GAGGGAAGAGCCATCAATGTC--TGCTCCAGAAA--CCTACTGGTGAAGAGGG--TT 1457
      756  GAGGGAAGAGCCATCAATGCTCTGCTCCAGGAAACCTACTTGGTGAAGAAGGGGTTT 815
      1458  TGGTGAAGGGGGCAAGAGGCC 1478
      816  GGTGAAGGGGGCAAGAGCCC 836

RESULT 12
LOCUS   BX459685
DEFINITION  BX459685 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
          CS0DF003YD04 5-PRIME, mRNA sequence.
ACCESSION  BX459685
VERSION    BX459685.1 GI:31035114
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS    Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by Life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 8786.r For
          more information about this cluster, see
          http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF003DB02QP1&cluster=8786.r. Contact :
          Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DF003DB02QP1.
          Location/Qualifiers
            1..972
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DF003YD04"
              /tissue_type="FETAL BRAIN"
              /dev_stage="fetal"
              /clone_lib="Homo sapiens FETAL BRAIN"
              /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
              was primed with a NotI-oligo(dT) primer. Five prime end
              enriched, double-strand cDNA was digested with Not I and
              cloned into the Not I and EcoRV sites of the pCMVSPORT 6
              vector. Library was not normalized."

ORIGIN
Query Match 47.7%; Score 722; DB 13; Length 972;
Best Local Similarity 97.2%; Pred. No. 5.7e-149;
Matches 754; Conservative 0; Mismatches 0; Indels 22; Gaps 1;

QY      1  ATGAGAGGGGGTCCAGTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
      197  ATGAGAGGGGGTCCAGTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 256
      61  GCGACCATCATGTGACTACTCTGGAGAGGCGAGATGGTGCCAGAGCCCTACTAGAAC 120
      257  GCAGCCATCATGTGACTACTCTGGAGAGGCGAGATGGTGCCAGAGCCCTACTAGAAC 316
      121  GGTGTGAGCCCCCACCACAGGGCCAGAGCTGCCTCTGTGATCCCTGSCAGTACTCAAGA 180
      317  GGTGTGAGCCCCCACCACAGGGCCAGAGCTGCCTCTGTGATCCCTGSCAGTACTCAAGA 376
      181  CTGCTCCCGCGCGCTAGCCTCTCAGCCAGGAAGAGCTTTCCTTACAGGAGCGGCCAGCA 240
```

```
Db      377  CTGCTCCAGCCCGCCTAGCCTCTCAGCCAGGAAGCTTTCCTTACAGAGCGGCCAGCA 436
      241  GGAAGCTATCTGGAGCGCAGGCTGGGCTTATGCAAGGGGCTTGCAGCCACACATCTCC 300
      437  GGAAGCTATCTGGAGCGCAGGCTGGGCTTATGCAAGGGGCTTGCAGCCACACATCTCC 496
      301  CCCCAGGCTTGGCGGAGGCCACCATCGAGTCCCCACCAAGCTGGCCCATCTCAGATGCAGAG 360
      497  CCCCAGGCTTGGCGGAGGCCACCATCGAGTCCCCACCAAGCTGGCCCATCTCAGATGCAGAG 556
      361  GACTCGGTGACCTGAAACAGTACAGCTGACAGTGAATGGCAAGGTCCTACGGT 420
      557  GACTCGGTGACCTGAAACAGTACAGCTGACAGTGAATGGCAAGGTCCTACGGT 616
      421  GTGGTGAGGCTGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 480
      617  GTGGTGAGGCTGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 676
      481  AAAAAAGAAAGTTACTGAAGCAGTATGGCTTTCACGCTCGCCCTCCCCCGAGAGGGTCCCAG 540
      677  AAAAAAGAAAGTTACTGAAGCAGTATGGCTTTCACGCTCGCCCTCCCCCGAGAGGGTCCCAG 736
      541  GTCGCCAGGAGGAGCCAGCAGAGCTGCTGCCCTGGAGGGGTGTACCAAGAGATT 600
      737  GTCGCCAGGAGGAGCCAGCAGAGCTGCTGCCCTGGAGGGGTGTACCAAGAGATT 796
      601  GCATCTCTGAAGAAGCTGGACCAAGCTGAATGTGGTCAAACTGATCGAGTCTCTGGAATGAC 660
      797  GCATCTCTGAAGAAGCTGGACCAAGCTGAATGTGGTCAAACTGATCGAGTCTCTGGAATGAC 856
      661  CCAGCTGAGGACCAACTTATTGTGTGTGACCTCTCTGAGAAAGG-----706
      857  CCAGCTGAGGACCAACTTATTGTGTGTGACCTCTCTGAGAAAGGCGGTGAGTTCCTCCGT 916
      707  -----GGCCGCTCATGGAAGTCCCTGTGACAGCCCTCTCGAGAGGCAAG 754
      917  CCTGATCAGGCGCGCTCATGGAAGTGCCTGTGACAGCCCTTCTCGGAGGAGCAAG 972

RESULT 13
LOCUS   BG396486
DEFINITION  602459214F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4581651 5',
          mRNA sequence.
ACCESSION  BG396486
VERSION    BG396486.1 GI:13289845
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS    NIH-MGC http://mgi.ncl.nih.gov/;
          National Institutes of Health,
          Unpublished (1999)
JOURNAL     Contact: Robert Strausberg, Ph.D.
COMMENT     Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
          Plate: LLCW1302 row: 1 column: 04
          High quality sequence stop: 733.
          Location/Qualifiers
            1..893
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4581651"
              /tissue_type="retinoblastoma"
```


Db 314 ATTTTCATCTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTTTCTGTGAGGAGC 373
QY 1127 CAGAAATCAGCGAGGAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGA 1186
Db 374 CAGAAATCAGCGAGGAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGA 433
QY 1187 CGAGAAATGGGGTGCAGACATCAAGTTGCACCTTGGGTGACCAAGAACGGGGAGGAGC 1246
Db 434 CGAGAAATGGGGTGCAGACATCAAGTTGCACCTTGGGTGACCAAGAACGGGGAGGAGC 493
QY 1247 CCCTTCCTTCGAGGAGAGGAGCTGACAGGCTGGTGAAGTGAACAGGGGGAGGTTAAGA 1306
Db 494 CCCTTCCTTCGAGGAGGAGGAGCTGACAGGCTGGTGAAGTGAACAGGGGGAGGTTAAGA 553
QY 1307 ACTCAGTCAGGCTCATCCCGAGCTGACACAGCGTGATCTGTGAAGTCCATGCTGAGGA 1366
Db 554 ACTCAGTCAGGCTCATCCCGAGCTGACACAGCGTGATCTGTGAAGTCCATGCTGAGGA 613
QY 1367 AGCGTTCTCTTTGGGAACCGCTTTGAGCCCCAGGCACGAGGAGGAGCGATCCATGTCG 1426
Db 614 AGCGTTCTCTTTGGGAACCGCTTTGAGCCCCAGGCACGAGGAGGAGCGATCCATGTCG 673
QY 1427 CTCACGAAACCTACTGTTGAAGAGGCTTTGGTGAAGGGGGGCAAGCCAGAGCTCC 1486
Db 674 CTCACGAAAC--TACTGTGGGAAGAGGGTTGTTGAAAGGGGGCAAGAG-CCAAAGCTCC 730
QY 1487 CCGGCGTCCAGGAAGACGAGGCTGCATCC 1515
Db 731 CCGGCGTTCAGGAAGACAGGCTGATCC 759

RESULT 15
LOCUS BQ571613
DEFINITION UI-W-FCO-byc-e-19-0-UI.r1 NIH_BMAP_FCO Mus musculus cDNA clone
IMAGE:5716218 5', mRNA sequence.
ACCESSION BQ571613
VERSION BQ571613.1 GI:21474930
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 784)
NIH-MGC <http://imgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
Location/Qualifiers
1..784
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5716218"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FCO"
/notes="Organ: brain; Vector: pYX-Asc; Site: 1: Ecor I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 44.1%; Score 668.6; DB 13; Length 784;
Best Local Similarity 91.6%; Pred. No. 3.4e-137;
Matches 718; Conservative 0; Mismatches 65; Indels 1; Gaps 1;
QY 261 GCGTGGGCTTATGCGACGGGCGCTGCGAGCCACATCTCCCCCGGGGCTGCGGGAGGCC 320
Db 2 GGTGGGCTTACTCTACAGGACCTGCCAGTCACATCTCTCTCGGTCTCTGGCGGAGACC 61
QY 321 CACCATCGAGTCCCAACCAACGTCGCATCTCAGATCAGAGGACTCGTGCAGCTGAACCA 380
Db 62 CACCATCGAGTCCCAACCGGTGTGGCCATCTCAGACAGAGGACTGTGTGCAACTGACCA 121
QY 381 GTACAAGCTGCGAGGTGAGATTGGCAAGGTCCTACGGTGTGGTGAAGCTGGCTTACAA 440
Db 122 GTACAAGCTGCGAGGTGAGATTGGCAAGGTCCTATGTTGGTGAAGCTGGCTTACAA 181
QY 441 CGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTACTAGACA 500
Db 182 CGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTACTAGACA 241
QY 501 GTATGGCTTTCCACGTGCGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACCAGC 560
Db 242 GTATGGCTTTCTCGCGGCTCTCCCGAGAGGGTCCCAAGCTACCCAGGGAGGCGCAGC 301
QY 561 CAAGCAGCTGCTGCCCTCGAGGCGGTGTACCAGGAGATTGGCATCTCTGAAGAGCTGGA 620
Db 302 CAACAGCTGCTGCCCTCGAGGCGGTGTGTATCAGGAGATTGCCATTTAAAGAGCTGA 361
QY 621 CCACGTGAATGTGGTCAAACTGATCGAGTCTCTGATGATCCAGTCCAGGACCACTCTA 680
Db 362 CCACGTGAATGTAGTCAAAATTCATCGAGTCTCTGGATGATCCCGCTGAAGACAACTCTA 421
QY 681 TTTGGTGTGTTGACCTCTCGAGAAAGGGGCCCGTCTATGGAAGTGCCTGTGACAAGCCCTT 740
Db 422 TTTGGTGTGTTGACCTCTCGAGAAAGGGGCCAGTCTATGGAAGTGCCTGTGACAAGCCCTT 481
QY 741 CTCCGAGGAGCAAGCTCGCTCTTACCTCGGGAACGTCATCTTGGGCTCGAGTACTTGA 800
Db 482 CCCAGAGGAGCAAGCTCGCTCTTACCTCGGGAACATCTCTGGGCTCGAGTACTTGA 541
QY 801 CTGCCAAGAGATGCTGCCACAGGGAATCAAGCCATCAACCTGCTCTCTGGGGATGATGG 860
Db 542 CTGCCAAGAGATTGTCCACAGGGAATCAAGCCATCTAACCTACTCTTGGGGACACCG 601
QY 861 GCACGTGAAGATCGCGACTTTGGGCTCAGCAACCAAGTTTGGGGGAACAGCCTCAGCT 920
Db 602 GCATGTGAAGATCGCGACTTTGGGCTCAGTAAACAGTTTGAAGGAATGATGCTCAGCT 661
QY 921 GTCCAGCAGCGGGGAACCCCGAGCAATTCATGGCCCCCGAGGGCAATTTCTGATTCGGCCA 980
Db 662 GTCCAGTACGGCAGGACCCCGCATTCATGGCCCCCGAGGGCAATTTCTGATTCGGGCA 721
QY 981 GAGCTTCAGTGGGAAGCGCTTGGATGTATGGGCCACTGGGTCACGTTGACTGCTTTGT 1040
Db 722 GAGCTTCAGTGGGAAGG-CTTGGATGTATGGGCCACTGGGTCACGCTATATGCTTTGT 780
1041 CTAT 1044
781 CTAT 784

Search completed: July 25, 2004, 11:34:37
Job time : 6003 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 06:22:55 ; Search time 140 Seconds
(without alignments)
6005.361 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggagggggtccagctgt.....aggagacgaggtgcattcc 1515

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	100.0	2190	US-09-729-995-1	Sequence 1, Appli
2	1515	100.0	2190	US-10-133-689-1	Sequence 1, Appli
3	361	23.8	29629	US-09-729-995-3	Sequence 3, Appli
4	361	23.8	29629	US-10-133-689-3	Sequence 3, Appli
5	142.8	9.4	288	US-09-016-434-1008	Sequence 1008, Ap
6	99.4	6.6	236	US-09-016-434-809	Sequence 809, Appl
7	90.4	6.0	1302	US-08-913-050A-2	Sequence 2, Appli
8	90.4	6.0	1302	US-09-016-434-1146	Sequence 1146, Ap
9	87	5.7	1466	US-08-749-902-2	Sequence 2, Appli
10	83.8	5.5	2968	US-08-685-852-1	Sequence 1, Appli
11	83.4	5.5	1599	US-09-256-465-1	Sequence 1, Appli
12	83.4	5.5	1599	US-09-167-322-3	Sequence 3, Appli
13	83.4	5.5	1599	US-09-023-655-1004	Sequence 1004, Ap
14	82.8	5.5	1257	US-09-793-875-15	Sequence 15, Appli
15	82.8	5.5	1826	US-09-799-875-13	Sequence 13, Appli
16	82.8	5.5	1864	US-09-819-607-1	Sequence 1, Appli
17	82.8	5.5	1975	US-08-852-743-1	Sequence 1, Appli
18	82.8	5.5	1975	US-09-185-370-1	Sequence 1, Appli
19	82.2	5.4	1636	US-09-016-434-1433	Sequence 1433, Ap
20	79.4	5.2	2158	US-09-765-815-1	Sequence 1, Appli
21	79.4	5.2	3604	US-09-688-188B-27	Sequence 27, Appli
22	79.4	5.2	3604	US-09-291-417D-27	Sequence 27, Appli
23	78.6	5.2	1940	US-09-718-032-1	Sequence 1, Appli
24	78.6	5.2	1941	US-09-082-737-1	Sequence 1, Appli
25	78.6	5.2	2050	US-09-688-188B-28	Sequence 28, Appli
26	78.6	5.2	2050	US-09-291-417D-28	Sequence 28, Appli
27	78.6	5.2	2806	US-09-688-188B-102	Sequence 102, Appl

ALIGNMENTS

RESULT 1

US-09-729-995-1
; Sequence 1, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Human
US-09-729-995-1
Query Match 100.0%; Score 1515; DB 4; Length 2190;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGGGGGGTCCAGCTGTCTGCTGCAGAGATCCTCGGGCAGAGCTGTTAGAACGGGTG	60
DB	173	ATGAGGGGGGTCCAGCTGTCTGCTGCAGAGATCCTCGGGCAGAGCTGTTAGAACGGGTG	232
QY	61	CGAGCATCGATGTGACTCCTCTGGAGGAGGAGATGTTGGCCAGAGCTCTACTAGAAAC	120
DB	233	CGAGCATCGATGTGACTCCTCTGGAGGAGGAGATGTTGGCCAGAGCTCTACTAGAAAC	292
QY	121	GCTGTGGACCCCAACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	180
DB	293	GCTGTGGACCCCAACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	352
QY	181	CTGCTCCCAAGCCCGCTTAGCTCTCAGCCAGAGAGCTTTCCTACAGAGGGCCAGCA	240
DB	353	CTGCTCCCAAGCCCGCTTAGCTCTCAGCCAGAGAGCTTTCCTACAGAGGGCCAGCA	412
QY	241	GGAAGCTATCTGGAGGCGCAGGCTGGGGCTTTATGACAGGGGCTTGCAGGCCACATCTCC	300
DB	413	GGAAGCTATCTGGAGGCGCAGGCTGGGGCTTTATGACAGGGGCTTGCAGGCCACATCTCC	472
QY	301	CCCCGGGCTTGGGGAGGCCCAACCATCGAGTCCCAACACGTGGCCATCTCAGATCAGAG	360
DB	473	CCCCGGGCTTGGGGAGGCCCAACCATCGAGTCCCAACACGTGGCCATCTCAGATCAGAG	532
QY	361	GACTCGTGCAGCTGAACCAAGTACAGCTGAGATGAGTGGCAAGGTGCTTACCGT	420

Db 533 GACTGCGCTGAGCTGAACCACTACAAGCTGCAGAGTGAGATTGSCAAGGGTGCCCTACGGT 592
Qy 421 GTGTGAGGTGGCTCAACAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC 480
Db 593 GTGTGAGGTGGCTCAACAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC 652
Qy 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGTCCTCAG 540
Db 653 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGTCCTCAG 712
Qy 541 GCTGCCAGGAGACCAAGCAGCTGCTGCCCTCTGGAGCGGGTGACAGAGATT 600
Db 713 GCTGCCAGGAGACCAAGCAGCTGCTGCCCTCTGGAGCGGGTGACAGAGATT 772
Qy 601 GCCATCTGAAGAAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGTCCTGGATGAC 660
Db 773 GCCATCTGAAGAAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGTCCTGGATGAC 832
Qy 661 CCAGCTGAGGACAACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGCCGGTCATGAA 720
Db 833 CCAGCTGAGGACAACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGCCGGTCATGAA 892
Qy 721 GTGCCCTGTGACAAGCCCTCTTCGGAGGAGCAAGCTCGCCTCTACCTGCGGGAGCTCATC 780
Db 893 GTGCCCTGTGACAAGCCCTCTTCGGAGGAGCAAGCTCGCCTCTACCTGCGGGAGCTCATC 952
Qy 781 CTGGGCTTCAGTACTTGCACGAGAGATCGTCCACAGGAGCATCAAGCCATCCAAC 840
Db 953 CTGGGCTTCAGTACTTGCACGAGAGATCGTCCACAGGAGCATCAAGCCATCCAAC 1012
Qy 841 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCCACTTTGGCGTCAGCAACCACTTT 900
Db 1013 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCCACTTTGGCGTCAGCAACCACTTT 1072
Qy 901 GAGGGGAACACGCTCAGCTGTTCAGCACCGCGGGAAACCCAGCAATTATGGCCCCCGAG 960
Db 1073 GAGGGGAACACGCTCAGCTGTTCAGCACCGCGGGAAACCCAGCAATTATGGCCCCCGAG 1132
Qy 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAGGCCCTTGATGTATGGGCCATGGC 1020
Db 1133 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAGGCCCTTGATGTATGGGCCATGGC 1192
Qy 1021 GTCACGTTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCCTCGCC 1080
Db 1193 GTCACGTTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCCTCGCC 1252
Qy 1081 CTCACAGGAAGATCAAGAAATGAGCCCGTGTTCCTCAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAAGATCAAGAAATGAGCCCGTGTTCCTCAGGAGCCAGAAATCAGCGAG 1312
Qy 1141 GAGCTCAAGGACCTGATCTCTGAGATGTTAGACAGAAATCCGAGAGGAAATGGGGTG 1200
Db 1313 GAGCTCAAGGACCTGATCTCTGAGATGTTAGACAGAAATCCGAGAGGAAATGGGGTG 1372
Qy 1201 CCAGACATCAAGTTGCAACCTTCGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1260
Db 1373 CCAGACATCAAGTTGCAACCTTCGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1432
Qy 1261 GAGGAGCACTGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1492
Qy 1321 ATCCCGAGCTGGACCAAGGTGATCTGGTCAAGTCCATGCTGAGGAGCGTCTCTTTGG 1380
Db 1493 ATCCCGAGCTGGACCAAGGTGATCTGGTCAAGTCCATGCTGAGGAGCGTCTCTTTGG 1552
Qy 1381 AACCCGTTTGAAGCCCAAGGCAAGAGGGAAGAGCGATCCATGCTCTCCAGGAAACCTA 1440
Db 1553 AACCCGTTTGAAGCCCAAGGCAAGAGGGAAGAGCGATCCATGCTCTCCAGGAAACCTA 1612
Qy 1441 CTGGTGAAGAAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGCTCCCCCGCTCCAGGAA 1500
Db 1613 CTGGTGAAGAAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGCTCCCCCGCTCCAGGAA 1672

Qy 1501 GACGAGCTGCATCC 1515
Db 1673 GACGAGCTGCATCC 1687

RESULT 2

US-10-135-689-1
; Sequence 1, Application US/10135689
; Patent No. 6670162
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-135-689-1

Query Match 100.0%; Score 1515; DB 4; Length 2190;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGGGGGTCCAGCTGCTGCCAGGATCTTCGGGCAGAGCTGGTAGAACGGGTG 60
Db 173 ATGAGGGGGGTCCAGCTGCTGCCAGGATCTTCGGGCAGAGCTGGTAGAACGGGTG 232
Qy 61 GCAGCATCATGTGACTCACTTGGAGGAGGAGAGTGGTGGCCAGAGCTTACTAGAAAC 120
Db 233 GCAGCATCATGTGACTCACTTGGAGGAGGAGAGTGGTGGCCAGAGCTTACTAGAAAC 292
Qy 121 GGTGTGAGCCCCCACACAGGGCCAGAGCTGCTCTGTGATCCTGGCAGTACTTCAGA 180
Db 293 GGTGTGAGCCCCCACACAGGGCCAGAGCTGCTCTGTGATCCTGGCAGTACTTCAGA 352
Qy 181 CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db 353 CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 412
Qy 241 GGAAGCTATCTGGAGGGGAGGCTGGGCTTATGCCAGGGGCTTCGACGCCCATCTCC 300
Db 413 GGAAGCTATCTGGAGGGGAGGCTGGGCTTATGCCAGGGGCTTCGACGCCCATCTCC 472
Qy 301 CCCCAGGCTGGCGGAGGCCACCATCTGAGTCCACACACGTGGCCATCTCAGATGCAAG 360
Db 473 CCCCAGGCTGGCGGAGGCCACCATCTGAGTCCACACACGTGGCCATCTCAGATGCAAG 532
Qy 361 GACTCGTGACGTGAACCAAGTCAAGCTGCAGGTGAGATTGGCAAGGGTGCCTACGGT 420
Db 533 GACTCGTGACGTGAACCAAGTCAAGCTGCAGGTGAGATTGGCAAGGGTGCCTACGGT 592
Qy 421 GTGGTCAGGCTGGCTTACCAACGAAAGTGAAGACAGACATATGCAATGAAGTCCCTTCC 480
Db 593 GTGGTCAGGCTGGCTTACCAACGAAAGTGAAGACAGACATATGCAATGAAGTCCCTTCC 652
Qy 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCCCGAGAGGGTCCCAG 540
Db 653 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCCCGAGAGGGTCCCAG 712
Qy 541 GCTGCCAGGAGGACCAAGCAGAGCTGTGCCCCCTGGAGCGGGTGTACCAAGGAGATT 600
Db 713 GCTGCCAGGAGGACCAAGCAGAGCTGTGCCCCCTGGAGCGGGTGTACCAAGGAGATT 772


```
QY 1 ATGAGGGGGTCCAGCTGTCTGTCAGAGATCCTCGGCGAGAGCTGGTAGAACGGGTG 60
Db |||||||
QY 3000 ATGAGGGGGTCCAGCTGTCTGTCAGAGATCCTCGGCGAGAGCTGGTAGAACGGGTG 3059
Db |||||||
QY 61 GCAGCCATCGATGTGACTCACTTGGAGAGGAGCAGATGTTGGCCAGAGCCTACTAGAAC 120
Db |||||||
QY 3060 GCAGCCATCGATGTGACTCACTTGGAGAGGAGCAGATGTTGGCCAGAGCCTACTAGAAC 3119
Db |||||||
QY 121 GGTGTGAGACCCCAACCAAGGCGCAGAGCTGCCTCTGTGATPCCCTGGCAGTACTCAAGA 180
Db |||||||
QY 3120 GGTGTGAGACCCCAACCAAGGCGCAGAGCTGCCTCTGTGATPCCCTGGCAGTACTCAAGA 3179
Db |||||||
QY 181 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTTCCCTACAGGAGCGCCAGCA 240
Db |||||||
QY 3180 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTTCCCTACAGGAGCGCCAGCA 3239
Db |||||||
QY 241 GGAAGCTATCTGAGAGGCGCAGGCTGGGCTTTATGCCACGGGGCTGCCAGCCACATCTCC 300
Db |||||||
QY 3240 GGAAGCTATCTGAGAGGCGCAGGCTGGGCTTTATGCCACGGGGCTGCCAGCCACATCTCC 3299
Db |||||||
QY 301 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACCAAGTGGGCATCTCAGATGCAGAG 360
Db |||||||
QY 3300 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACCAAGTGGGCATCTCAGATGCAGAG 3359
Db |||||||
QY 361 G 361
Db |||||||
QY 3360 G 3360
Db |||||||
```

RESULT 5

```
US-09-016-434-1008
; Sequence 1008, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1008:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 926034
; US-09-016-434-1008
```

```
Query Match 9.4%; Score 142.8; DB 4; Length 288;
Best Local Similarity 70.7%; Pred. No. 1.7e-27;
Matches 203; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

QY 472 GTCTTTTCCAAAAGAAAGTTACTGAAGCAGTATGGCTTTTCCACGTGCGCTTCCGCCGAGA 531
Db |||||||
QY 532 GGGTCCAGGTCGCCAGGAGGAGCAGCCAGCAGCTGCTGCCCTTGGAGCGGGTGTAC 591
Db |||||||
QY 61 GGCACCCCGGCAGCTCCTGGAGGCTGCATCAGCCCGAGGGGCCCATTTGAGCAGGTGTAC 120
Db |||||||
QY 592 CAGGAGATTGCATCTCTGAAGAGCTGCACCATGATGTGGTCAAACTGATCGAGTGC 651
Db |||||||
QY 121 CAGGAATTGCCATCTCAAGAGCTGACCCCATTTGGTGAAGCTGGTGGAGGTC 180
Db |||||||
QY 652 CTGATGACCCAGCTGAGGACCACTCTATTTGGTGTGTTGACCTCTGAGAAAGGGGCC 711
Db |||||||
QY 181 CTGATGACCCCAATGAGGACCATCTGTACATGGTGTTCGA-ATGGTCAACCAAGGGCCC 239
Db |||||||
QY 712 GTCATGGAAGTGGCTGTGACAGCCCTTCTCGAGGAGGAGCAAGCTCG 758
Db |||||||
QY 240 GTGATGGAAGTGCCCACTTAAACCCCTTTTGAAGAGCAGGGCCC 286
Db |||||||
```

RESULT 6

```
US-09-016-434-809
; Sequence 809, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 809:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

IMMEDIATE SOURCE:
LIBRARY: BRAINOT03
CLONE: 531037
US-09-016-434-809
Query Match 6.6%; Score 99.4; DB 4; Length 236;
Best Local Similarity 71.8%; Pred. No. 2.3e-16;
Matches 155; Conservative 0; Mismatches 59; Indels 2; Gaps 2;
QY 703 AAGGGCCGCTCATGAAAGTGGCC-TGTGACAAAGCCCTTCTCGAGAGCAAGCTCGCT 761
DB 20 AAGGGCCGCTGATGAAAGTGGCCCTTAAACCACTCTCTGAAGACCAGGCGCTT 79
QY 762 CTACCTGC-GGAGCTCATCTCTGGGCTCGAGTACTTGCACCTGCCAGAGATCGTCCACA 820
DB 80 CTACTTCAGGGATCTGATCAAGGATCGAGTACTTACACTACAGAGATCATCCACC 139
QY 821 GGGACATCAAGCCATCCAACTCTCTCTGGGGGATGATGGGCACTGGAAGATCGCGACT 880
DB 140 GTGACATCAAACTTTCAACTCTCTGCTCGAGAGATGGGACATCAAGATCGCTGACT 199
QY 881 TTGGCTCAGCAACCAAGTTTGGGGGACGACGCTC 916
DB 200 TTGGTGTGAGCAATGAATTCAGGGGACGTGACGCGC 235

RESULT 7
US-08-913-050A-2
Sequence 2, Application US/08913050A
Patent No. 5827726
GENERAL INFORMATION:
APPLICANT: NEZU, Jun-ichi
TITLE OF INVENTION: DNA ENCODING PROTEIN KINASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 7th Street N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,050A
FILING DATE: 05-SEP-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 57104/1995
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP PCT/JP96/00660
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: NEZU=4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-913-050A-2
Query Match 6.0%; Score 90.4; DB 1; Length 1302;
Best Local Similarity 50.6%; Pred. No. 9.7e-14;

Matches 335; Conservative 0; Mismatches 306; Indels 21; Gaps 4;
QY 586 GTGTACACAGGAGATTGCCATCCTGAAGAAGCTGGACCACTGAATGTGTGTTCAAACTGATC 645
DB 283 GTGAAGAAGGAAATTCAACTACTGAGGAGTTACGGCACAAAATGTATCATCCAGCTGGTG 342
QY 646 GAGTCTCTGGATGACCCAGCTGAGGACAACTCTATTTGGTGTGTTGA-----CCTCCTG 699
DB 343 GATGTGTTATACACGAAGAAGACAGAAAATGTATGTTGATGAGTACTGCTGTGT 402
QY 700 AGAAGGGGCGCTCATGGAAGTGCCTGTGACAGCCCTTCTCGAGAGCAAGCTCGC 759
DB 403 GGCATGCAAGGAAATGCTGGACAGCTGCCGAGAGACGTTTCCAGTGTGCCAGGCCAC 462
QY 760 CTCTACCTGCGGAGCTCATCTGGGCTCGAGTACTTGCACCTGCCAGAAATCGTCCAC 819
DB 463 GGGTACTTCTGTGACTGATTGACGGCTGGAGTACTTGCATAGCCAGGCACTTGTGCAC 522
QY 820 AGGACATCAAGCCATCAAACTCTCTCTGGGGGATGATGGGCACTGGAAGTTCGCCGAC 879
DB 523 AAGGACATCAAGCCGGGAACTCTGTGCTTACACCGGTGGCACCTCTCAAAATCTCCGAC 582
QY 880 TTTGGCTCAGCAACCAAGTTTGGGGGACGACGCTCAGCTGTCCAGCACGCG-----G 933
DB 583 CTGGGCTGCGGCGAGGCACTGCCCGCTTCGGGCGGACGACCTGCCGAGCCAGCG 842
QY 934 GGAACCCCAAGATTTCATGGCCCGGAGCCATTTCTGATTCGGGCGGACAGCTTTCAGTGG 993
DB 643 GGTCTCCCGGCTTTCAGCGCGCCGAG--ATTGCCAAGCGGCTGGACACCTTCTCCGCG 699
QY 994 AAGGCTTGGATGATGAGGCACTGGGCTCAGTGTGACTGTCTATGGGAAGTGC 1053
DB 700 TTCAAGGTGGACATCTGTGCTGGCTGGGCTCACCTCTACACATCACCAAGGCTCTGTAC 759
QY 1054 CCATTTCAGGAGATTTCATCTGCGCTCCACAGAGATCAAGATGAGCCGCTGGTG 1113
DB 760 CCGTTCGAAGGGGACAACTACAGTTGTTGAGAACATCGGAAGGGAGCTACGCC 819
QY 1114 TTTCTCTGAGGAGCCAGAAATCAGGAGGCTCAAGGACCTGATCCTGAAGATGTTAGAC 1173
DB 820 ATCCCGGGGACTGTG-----GCCCGCGCTCTCTGACCTCTGAAAGGATGCTTGAG 873
QY 1174 AAGAATCCGAGACGAGAAATTTGGGTGCCAGACATCAAGTTCGACCTTGGGTGACCAAG 1233
DB 874 TAGGAACCGGCCAAGAGGTTCTCCATCCGCGAGATCCGCGACGACAGCTGTTCCGGAAG 933
QY 1234 AA 1235
DB 934 AA 935

RESULT 8
US-09-016-434-1146
Sequence 1146, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1146:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91480860
; US-09-016-434-1146

Query Match 6.0%; Score 90.4; DB 4; Length 1302;
Best Local Similarity 50.6%; Pred. No. 9.7e-14;
Matches 335; Conservative 0; Mismatches 306; Indels 21; Gaps 4;

Qy 586 GTGTACCAGGAGATTGCCATCCTGAAGAAGCTGGACCACTGTAATGTGGTCAAACTGATC 645
Db |||||
Qy 283 GTGAAGAAGGAAATTCAACTACTGAGGAGTTACGGCAGAAATGTATCCAGCTGGTG 342
Db |||||
Qy 646 GAGCTCTGTGATGACCACTGAGACAACTCTATTGGTGTGTTGA-----CCTCTG 699
Db |||||
Qy 343 GATGTGTATACAAAGAGAGAGAGAGAAATGTATGTGTGATGGAGTACTGGTGTGT 402
Db |||||
Qy 700 AGAAGGGGGCCGTCATGGAGTGCCCTGTGACAGCCCTCTCGGAGGACAGCTCGC 759
Db |||||
Qy 403 GGCATGACGAGAAATGTGGACAGGTGCGGAGAGCGTTTCCAGGTGTGCGAGCCAC 462
Db |||||
Qy 760 CTCTACTCGGGAGCTGATCCTGGGCTCGAGTACTTGACCTCCAGAGATCGTCCAC 819
Db |||||
Qy 463 GGGTACTTCTGTACGTGATGACGGCTGGAGTACCTGCATAGCCAGGGCAATTTGCAC 522
Db |||||
Qy 820 AGGACATCAAGCATCCAACTGCTCTGGGGATGATGGGACAGTGAATGCCGAC 879
Db |||||
Qy 523 AAGGACATCAAGCCGGGAACTCTGTCTCACCACCGGTGGCACCCCTCAAAATCTCGAC 582
Db |||||
Qy 880 TTGGCGCTCAGCAACCAAGTTTGGGGGAACAGACCTCAGCTGTCCAGCAGGC-----G 933
Db |||||
Qy 583 CTGGCGTGGCCGAGGCACTGCACCGTTTCGGCGGAGACACACTGCGGACAGCCAG 642
Db |||||
Qy 934 GGAACCCAGCATTCATGGCCCGGAGGCAATTTCTGATTCGGCCAGAGCTTCAGTGGG 993
Db |||||
Qy 643 GGCTCCCGCGCTTTCAGCGCCCGCGAG---ATTGCCAAAGCGCTGGACACCTTCTCGGC 699
Db |||||
Qy 994 AAGGCTTGTGATGTATGGGCACTGGCGTCAAGTTGTACTGCTTGTCTATGGAAGTGC 1053
Db |||||
Qy 700 TTCAAGTGGACATCTGCTGGGTGGGTGACCTCTACAAACATCACACGGGTCTGTAC 759
Db |||||
Qy 1054 CCATTTCATCGAGATTTCATCTTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG 1113
Db |||||
Qy 760 CCCTTCGAAGGGGACAACTCTACAAGTTGTTTGAAGACATCGGAAGGGGAGCTAGCC 819
Db |||||
Qy 1114 TTTCCTGAGGAGCCAGAAATCAGGAGGAGCTCAAGGACTGATPCTTGAAATGTTAGAC 1173
Db |||||
Qy 820 ATCCCGGGGAGTGTG-----GCCCGCCGCTCTCTGACCTGTGAAAGGGATGCTTGAG 873
Db |||||
Qy 1174 AAGAAATCCCGAGACGAGAAATTTGGGGTGGCCAGACATCAAGTTGACCCCTGGGTGACCAAG 1233
Db |||||
Qy 874 TAGCAACCGGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGCAGCTGGTTCGGGAAG 933
Db |||||
```

```
Qy 1234 AA 1235
Db ||
Db 934 AA 935

RESULT 9
US-08-749-902-2
; Sequence 2, Application US/08749902
; Patent No. 5985635
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,902
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0150 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1466 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: Consensus
; US-08-749-902-2
```

```
Query Match 5.7%; Score 87; DB 2; Length 1466;
Best Local Similarity 50.0%; Pred. No. 7.7e-13;
Matches 331; Conservative 0; Mismatches 310; Indels 21; Gaps 4;

Qy 586 GTGTACCAGGAGATTGCCATCCTGAAGAAGCTGGACCACTGTAATGTGGTCAAACTGATC 645
Db |||||
Qy 382 GTGAAGAAGGAAATCAACTACTGAGGAGTTACGGCAGAAATGTATCCAGCTGGTG 441
Db |||||
Qy 646 GAGGTCTCTGGATGACCCAGCTGAGGACAACTCTATTGGTGTGTTGA-----CCTCTG 699
Db |||||
Qy 442 GATGTGTATACACGAGAGAGAGAGAAATGTATGTGTATGGAGTACTGGTGTGT 501
Db |||||
Qy 700 AGAAGGGGCCGCTCATGGAAGTGCCTGTGAAGCCCTTCTCGGAGGAGCAAGCTCGC 759
Db |||||
Qy 502 GGCATGAGGAATGTGGACACGCTGCCGAGAGCGTTCCAGTGTGCCAGGCCAC 561
Db |||||
Qy 760 CTCTACTCGGGACCTCATCTTGGCCCTGAGTACTTGCACCTCCAGAGATCGTCCAC 819
Db |||||
Qy 562 GGGTACTTCTGTGCTGATGATGACGGCTGGGGTACTCTGATNGCCAGGNNATTTGTGAC 621
Db |||||
Qy 820 AGGACATCAAGCCATCCCACTCTCTCTGGGGGATGATGGGCACTGGGCACTGGAAGATCGCCGAC 879
Db |||||
```

Db 622 AAGGCATCAAGCCGGGAACCTGCTGCTCACCACCGGTGGCACCTCAAAATCTCCGAC 681
QY 880 TTGCGCTCAGCAACGAGTTTGGGGACGAGCTCAGCTGTCCAGCACGCG-----G 933
Db 682 CTGGGCGTGGCGAGGCACGTGACCCGTTTCGNCGGGACGACACCTGCGGACCGCCAG 741
QY 934 GGAAACCCAGCATTTCATGCCCCCGAGGCCATTTCTGATTCGGGCAGAGCTTCAGTGGG 993
Db 742 GACTCCCCGCTTTCAGCCGCCGA--NATGCAACGGCCTGGACACCTTCTCCGGC 798
QY 994 AAGGCTTGAGTATGAGGCCACATGGCGTCACTGTTGCTATGTTGTTATGAGGAGTGC 1053
Db 799 TTCAAGGTGGACATCTGGTCGGTGGGTTCACCTCTACACATCACCACGGGTCTGTAC 858
QY 1054 CCATTCATCGAGATTTCATCTGCGCTCCACAGAGATCAAGATGAGCCGTTGTFG 1113
Db 859 CCCTTCGAAGGGGACACATCTACAGTTGTTTGAGAACATCGGAGGGGAGCTACGCC 918
QY 1114 TTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC 1173
Db 919 ATCCCGGGGACTGTG-----GCCCGCGCTCTCTGACTCTGTAAGGGATGCTTGAG 972
QY 1174 AGAATCCGAGACGAGAAATGGGTTGCCAGACATCAAGTTGACCCCTTGGGTGACCAAG 1233
Db 973 TAGMAACCGGCAAGAGTTCTCCATCCGGCAGATCCGGCAGACACAGCTGGTTCCGGAAG 1032
QY 1234 AA 1235
Db 1033 AA 1034

RESULT 10
US-08-685-852-1
; Sequence 1, Application US/08685852
; Patent No. 6660837
; GENERAL INFORMATION:
; APPLICANT: KAIBUCHI, KOZO
; APPLICANT: ONO, YOSHITAKA
; APPLICANT: IWAMATSU, AKIHIRO
; TITLE OF INVENTION: MODIFIED PROTEIN DERIVED FROM PROTEIN KINASE N
; FILE REFERENCE: 016987/0844
; CURRENT APPLICATION NUMBER: US/08/685,852
; PRIOR FILING DATE: 1996-07-24
; PRIOR APPLICATION NUMBER: JP 7-262552
; PRIOR FILING DATE: 1995-09-14
; PRIOR APPLICATION NUMBER: JP 7-344606
; PRIOR FILING DATE: 1995-12-05
; PRIOR APPLICATION NUMBER: JP 8-080549
; PRIOR FILING DATE: 1996-03-08
; PRIOR APPLICATION NUMBER: JP 8-114226
; PRIOR FILING DATE: 1996-04-11
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)..(2862)
US-08-685-852-1

Query Match 5.5%; Score 83.8; DB 4; Length 2968;
Best Local Similarity 54.8%; Pred. No. 6.9e-12;
Matches 166; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 669 GGACACCTCTATTGGTGTTCACCTCTCTGAGAAAGGGCCCGCTCATGGAAGTGCCTG 728
Db 2100 GGAAGCAGTGTCTTCGTGATGAGTACTCGGCCGTGGGACCTGATGTGTCACATCCA 2159
QY 729 TGAACAGCCCTTCTCGGAGAGCAAGCTCGCTCTACCTCGGGAAGTCAATCTCGGCTT 788

Db 2160 CAGCGACGTGTTCTCTGAGCCCGGTGCCATCTTTTATTCGCTGCTGGTGTCTGGGCT 2219
QY 789 CGAGTACTTGCACTGCCAGAGATCGTCCACAGGACATCAAGCCATCCAACTGCTCTCT 848
Db 2220 ACAGTTTCTTACGAAACACAGATCGTCTACAGGACCTGAACTTGGACAATTTGCTCT 2279
QY 849 GGGGATGATGGGACGTGAAGATCGCGACTTTGGCGTCAAGAACCAAGTTTGGAGGAA 908
Db 2280 GGACACCGAGGGCTACGTCAAGATCGCAGACTTTGGCCTCTGCAAGGAGGGATGGGCTA 2339
QY 909 CGACCTCAGCTGTCCAGACGGCGGGAACCCAGCATTCATGGCCCGGAGGCCATTTC 968
Db 2340 TGGGACCGGACCGACCATTCGTGGGACCCCGAGTTCTTGGCCCTTGAGTGTCTGAC 2399
QY 969 TGA 971
Db 2400 GGA 2402

RESULT 11
US-09-256-465-1
; Sequence 1, Application US/09256465
; Patent No. 6043090
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-2 EXPRESSION
; FILE REFERENCE: RTS-0035
; CURRENT APPLICATION NUMBER: US/09/256,465
; CURRENT FILING DATE: 1999-02-23
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(1533)
US-09-256-465-1

Query Match 5.5%; Score 83.4; DB 3; Length 1599;
Best Local Similarity 51.3%; Pred. No. 6.7e-12;
Matches 254; Conservative 0; Mismatches 226; Indels 15; Gaps 2;
QY 706 GGGCCCGTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCTCTAC 765
Db 793 GAGCTGTTCTCCACCTGTCCGGGAGCGTGTCTCACAGAGGAGCGGCGCTTTAT 852
QY 766 CTGCGGAGCTCATCTGGGCTCTGAGTACTTGCATCTTGCATCTGCAAGATCTCCACAGGAC 825
Db 853 GGTGCAGAGATTGTCTCGGCTCTTGCATCTGCACTCGCGGACGTTGTATACCGGAC 912
QY 826 ATCAGCCATCCAACTGCTCTGGGGATGATGGCAGTGAAGATCCCGACTTTGC 885
Db 913 ATCAAGCTGGAACCTCATCTGGAAGATGGCCATCAAGATCATGACTGCTTGGC 972
QY 886 GTCAACACAGTTTGGGGGAACGCTCAGCTGTCTCAGCACGGCGGGAACCCAGCA 945
Db 973 CTCTGCAAGAGGGATCAGTACGGGGCCACCATGAACCTTCTGTGGAGCCCGGAG 1032
QY 946 TTATGGCCCCGAGGCCATTTCTGATTCGGCCAGAGCTTCACTGGGAAGGCTTGGAT 1005
Db 1033 TACTTGGCGCTGAGGTGTCTGGAGCAATGACTA-----TGGCCGGGCGGTGGAC 1083
QY 1006 GTATGGCCACTGGCGTCACTGTGCTTGTCTATGGAAGTGGCCATTCATCGAC 1065
Db 1084 TGGTGGGGCTGGGTGTGTCATGACGAGATGATGTGCGGCCCTGCTCCCTTCTACAC 1143
QY 1066 GATTTTCATCTCGGCCCTCCACAGGAAGATCAAGATGAGCCCGCTGTGTCTCTCAGGAG 1125
Db 1144 CAGGACAGAGGCGCTCTTCGAGCTCATCTCATGGAAGATCGCTTCCC-----G 1197
QY 1126 CCAGAAATCAGCAGGAGGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGATCCGAG 1185

Db 1198 CGCAGCTCAGCCCGAGCCCAAGTCCCTGCTGCTGGGCTGCTTAAGAGAGACCCCAAG 1257
QY 1186 ACGAGAATTGGGTG 1200
Db 1258 CAGAGGCTTGGTGG 1272

RESULT 12

US-09-167-322-3
; Sequence 3, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-167-322-3

Query Match 5.5%; Score 83.4; DB 4; Length 1599;
Best Local Similarity 51.3%; Pred. No. 6.7e-12;
Matches 254; Conservative 0; Mismatches 226; Indels 15; Gaps 2;
QY 706 GGGCCCGCTCATGGAGTGCCGTGTGACAGCCCTTCGGAGGAGCAAGCTCGCCTCTAC 765
Db 793 GAGCTGTCTTCCACTGTCCGGGAGCGGTCTTTCACAGAGGAGCGGCCCGGTTTAT 852
QY 766 CTGGGGGACGTCACTCTGGGCGCTCGAGTACTTGCACTGCCAGAGATCGTCCACAGGGAC 825
Db 853 GGTGCAGAGATTGCTCGGCTCTTGAGTACTTGCACTCGGGGACGTGGTATACCGGAC 912
QY 826 ATCAAGGCATCCAACTGCTCTGGGGGATGATGGGCACTGAGATCGCGACTTTGGC 885
Db 913 ATCAAGCTGGAACCTCATGCTGGACAAAGATGGCCACATCAAGATCACTGACTTTGGC 972
QY 886 GTCAAGCAACCAAGTTTGAGGGGAGCAAGCTGCTCCAGCAGCGGGGAAACCCAGCA 945
Db 973 CTCTGCAAGAGGGGCATCAGTGACGGGGCCACCATGAACCTTCTGTGGGACCCCGAG 1032

QY 946 TTCTATGGCCCGAGGCGCATTTCTGATTCCGCCAGAGCTTCAGTGGGAGGCGCTTGGAT 1005
Db 1033 TACCTGGCGCCTGAGGTGCTGAGGACAATGACTA-----TGGCGGGCGCGTGGAC 1083
QY 1006 GTATGGGCGCACTGGCGTCACGTTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGAC 1065
Db 1084 TGGTGGGGGCTGGGTGTGGTCACTGACGAGATGATGTCGGCGCGCTGCCCTTCTACAAC 1143
QY 1066 GATTTTCATCTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTGAGGAG 1125
Db 1144 CAGGACCACGAGCGCCTCTTCGAGCTCATCTCATGGAAGAGATCCGCTTCCC-----G 1197
QY 1126 CCAGAAATCAGGAGGAGCTCAAGACCTGATCCTGAGATGTTAGACAAGATCCCGAG 1185
Db 1198 CGCAGGCTCAGCCCGAGGCCAAGTCCCTGCTTGTGGGTGCTTTAAGAGAGACCCCAAG 1257
QY 1186 ACGAGAATTGGGTG 1200
Db 1258 CAGAGGCTTGGTGG 1272

RESULT 13

US-09-023-655-1004
; Sequence 1004, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1004:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g178325
; US-09-023-655-1004

Query Match 5.5%; Score 83.4; DB 4; Length 1599;
Best Local Similarity 51.3%; Pred. No. 6.7e-12;
Matches 254; Conservative 0; Mismatches 226; Indels 15; Gaps 2;

```

QY 706 GGGCCCGTCATGGAAGTGCCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 793 GAGCTGTTCTTCCACTGTCCGGAGCGGTCTTTCACAGAGAGCGGCGCGTTTAT 852
QY 766 CTGCGGAGCTCATCCTCGGCGCTCGAGTACTTGCACTGCCAGAAAGATCGTCCACAGGAC 825
DB 853 GGTGCAGAGATTGTCTCGGCTCTTGAGTACTTGCACTCGCGGAGCGTGTATACCGGAC 912
QY 826 ATCAAGCCATCAACTGCTCTCGGGGATGATGGGCAAGTGAAGATGCCGACTTTGGC 885
DB 913 ATCAAGCTGGAACAACCTCATGCTGACAAAGATGGCCACATCAAGATCACTGACTTTGC 972
QY 886 GTCAAGCAACAGTTTGGAGGGAACGACGCTCAGCTGTCCAGCACGCGGGAACCCAGCA 945
DB 973 CTCTGMAAGAGGATCATGTCACCGGGCCACCATGAAACCTTCTGTGGACCCCGAG 1032
QY 946 TTTCATGGCCCCCGAGGCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGAT 1005
DB 1033 TACCTGGGCGCTGAGTGTGGTGTACGAGATGATGTGCGGCGGCTGCCCTTCTACAC 1143
QY 1006 GTATGGGCCACTGSGCTCAGCTGTGCTTGTCTATGGGAAGTGGCCATTCATCGAC 1065
DB 1084 TGGTGGGGCTGGGTGTGGTGTACGAGATGATGTGCGGCGGCTGCCCTTCTACAC 1143
QY 1066 GATTTTCATCCTGGCCCTCCACAGGAAGATCAAGAATGAGCCGCTGTGTTCTCTGAGGAG 1125
DB 1144 CAGGACCAGAGCGCTCTCTCGAGCTCATCTCATGGAAGATCGGCTTCCC-----G 1197
QY 1126 CCAGAAATCAGGAGAGCTCAAGACCTGATCCTGAAGATGTTAGACAGAAATCCCGAG 1185
DB 1198 CGCAGCTCAGCCCGAGGCAAGTCCCTGCTTGTGGCTGCTTAAGAAGGACCCCAAG 1257
QY 1186 ACAGAAATGGGGTG 1200
DB 1258 CAGAGCTTGGTGGG 1272

RESULT 14
US-09-799-875-15
; Sequence 15, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17) ... (1273)
US-09-799-875-15

Query Match 5.5%; Score 82.8; DB 4; Length 1257;
Best Local Similarity 53.1%; Pred. No. 8.7e-12;
Matches 224; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY 586 GTGTACAGGAGATTGCCATCTCTGAAGAAGCTGGACCAAGTGGTCAAACTGATC 645
DB 211 GTCTTCGGGAGCTGGAGATCCTGAGGAGATCGACACGCTCTCTCGTGAACCTCTGG 270
QY 646 GAGGTCTTGATGACCCAGCTGAGGACAACTCTATTGTTGTTTTCACCTCTCGAAGAG 705

Query Match 5.5%; Score 82.8; DB 4; Length 1257;
Best Local Similarity 53.1%; Pred. No. 8.7e-12;
Matches 224; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY 586 GTGTACAGGAGATTGCCATCTCTGAAGAAGCTGGACCAAGTGGTCAAACTGATC 645
DB 211 GTCTTCGGGAGCTGGAGATCCTGAGGAGATCGACACGCTCTCTCGTGAACCTCTGG 270
QY 646 GAGGTCTTGATGACCCAGCTGAGGACAACTCTATTGTTGTTTTCACCTCTCGAAGAG 705
```

```

DB 271 TACTCTTCCAGGAC--GAGGAGGACATGTTATGTTGTTGGACCTGCTACTGCGGCGG 327
QY 706 GGGCCCGTCATGGAAGTGCCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 328 GACTGCGCTACCACCTGACGAGAACGTCAGTTCCTCGAGGACACGCTGAGGCTGTAC 387
QY 766 CTGCGGAGCTCATCCTCGGCGCTCGAGTACTTGCACTGCCAGAAAGATCGTCCACAGGAC 825
DB 388 ATCTCGAGATGGCACTGGCTCTGAGTACTGCTGCGCGGCCAGCATCATCCACAGAGAT 447
QY 826 ATCAAGCCATCAACTGCTCTCGGGGATGATGGGCAAGTGAAGATGCCGACTTTGGC 885
DB 448 GTCAAGCCCTGCAACATTTCTCTGATGAGAGAGACATGACACCTGACCGACTTCAAC 507
QY 886 GTCAAGCAACAGTTTGGAGGGAACGACGCTCAGCTGTCCAGCACGCGGGAACCCAGCA 945
DB 508 ATTGCCACATCATCAAGG---ACGGGAGCGGGCAGCGGCATTAGCAGGACCAAGCGG 564
QY 946 TTTCATGGCCCCCGAGGCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGAT 1005
DB 565 TACATGGCTCCGAGATCTTCCACTCTTTTGTCAACGCGCGGACCGGCTACTCTCTCGAG 624
QY 1006 GT 1007
DB 625 GT 626

RESULT 15
US-09-799-875-13
; Sequence 13, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17) ... (1273)
US-09-799-875-13

Query Match 5.5%; Score 82.8; DB 4; Length 1826;
Best Local Similarity 53.1%; Pred. No. 1e-11;
Matches 224; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY 586 GTGTACAGGAGATTGCCATCTCTGAAGAAGCTGGACCAAGTGGTCAAACTGATC 645
DB 227 GTCTTCGGGAGCTGGAGATCCTCGAGGAGATCGAGCACGCTTCTCTGTGAACCTCTGG 286
QY 646 GAGGTCTTGATGACCCAGCTGAGGACAACTCTATTGTTGTTTTCACCTCTCGAAGAG 705
DB 287 TACTCTTCCAGGAC---GAGGAGACATGTTTCATGTTGTTGTTGTTTTCACCTCTCGAAGAG 705
QY 706 GGGCCCGCTCATGGAAGTGCCTGTGACAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 344 GACTGCGCTACCACTCGCAGCAGAACGTCGAGTTCTCCGAGGACACGCTGAGGCTGTAC 403
QY 766 CTGCGGAGCTCATCTCTGGGCGCTGAGTACTGCACTGCGCAAGATCTCCACAGGAC 825
```



```

Db 404 ATCTGGAGATGGCACTGGCTCTGGACTACCTGGCGGCCAGCACATCATCCACAGAGAT 463
QY 826 ATCAAGCCATCCAACTCTCTCTGGGGGATGATGGGCAAGTGAAGATCGCGACTTTGGC 885
Db 464 GTCAAGCCCTGACAAATTTCTCTGGATGAGAGAGGACATGCACACCTGACCGACTTCAAC 523
QY 886 GTCAGCAACCAAGTTTGAAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCA 945
Db 524 ATTGCCACCATCATCAAG---ACGGGAGCGGGCGACGGCATTAGCAGGCAACCAAGCCG 580
QY 946 TTCATGGCCCCCGAGGGCCATTTCTGAATTCGGGCCAGAGCTTCAAGTGGGAAGGCCCTTGGAT 1005
Db 581 TACATGGCTCCGGAGATCTTCCACTCTTTTGTCAACGGGGGACCGGCTACTCCTTCGAG 640
QY 1006 GT 1007
Db 641 GT 642

```

Search completed: July 25, 2004, 11:37:16
 Job time : 147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 22:18:04 ; Search time 897 seconds
(without alignments)
7175.050 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687
Perfect score: 1515
Sequence: 1 atggagggggtccagctgt.....aggaagacgaggtgcattcc 1515

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	100.0	2190	6	ABL58698 Human kin
2	1510.2	99.7	1937	6	AAD30566 Human kin
3	1507	99.5	2018	6	ABZ11551 Human pol
4	1505.4	99.4	1611	6	ABX97035 Human NOV
5	1490.6	98.4	2711	6	ABK49563 Human CDN
6	1458.8	96.3	1542	4	AAS06710 Polynucle
7	1457.2	96.2	3501	5	AAD08635 Human kin
8	1455.8	96.1	1547	6	ABX97034 Human NOV
9	1381.4	91.2	1725	6	ABX97036 Human NOV
10	1236.6	81.6	3411	9	ADB53308 Primary r
11	660.2	43.6	2545	3	AAB229223 Human cel
12	651.2	43.0	4942	9	ADB75227 Prostate
13	651	43.0	1804	6	ABK92170 Prostate
14	651	43.0	1804	9	ADB75229 Prostate
15	643	42.4	2056	4	AAI57819 Human pol
16	641.6	42.3	4982	7	ABX34704 Human mdd
17	641.4	42.3	1767	4	AAF44686 Novel pro
18	585.6	38.7	2268	4	AAI59606 Human pol
19	585.6	38.7	2268	4	AAI59605 Human pol
20	529.6	35.0	705	4	AAS27190 cDNA enco
21	529.6	35.0	705	4	ABK43592 DNA enco
22	529.6	35.0	705	9	ADB93368 Human CDN
23	529	34.9	731	4	AAH99185 Human pro

24	526	34.7	2411	9	ADC99145 Human KPP
25	518	34.2	1927	4	AAI57820 Human pol
26	361	23.8	29629	6	ABL58699 Human kin
27	354.4	23.4	432	9	ADE85632 Farnesyl
28	348.8	23.0	439	4	AAH99412 Human pro
29	272	18.0	876	6	ABK34394 Human CDN
30	262.6	17.3	491	8	ACH14059 Human adu
31	244.6	16.1	485	8	ACH15568 Human adu
32	240	15.8	331	3	AAC02673 Human sec
33	202.4	13.4	1174	4	ABL19199 Drosophil
34	201.2	13.3	469	4	AAI15690 Probe #56
35	201.2	13.3	469	4	ABA57768 Human foe
36	201.2	13.3	469	4	AAI37350 Probe #60
37	201.2	13.3	469	4	ABA27140 Probe #56
38	201.2	13.3	469	4	AAK31464 Human bon
39	201.2	13.3	469	4	AAK05833 Human bra
40	201.2	13.3	469	4	ABK31146 Human liv
41	201.2	13.3	469	6	ABK06218 Human gen
42	201.2	13.3	10800	4	AAK83329 Human imm
43	201	13.3	218	4	AAI24914 Probe #14
44	201	13.3	218	4	ABA70385 Human foe
45	201	13.3	218	4	AAI50533 Probe #19

ALIGNMENTS

RESULT 1
ABL58698
ID ABL58698 standard; cDNA; 2190 BP.
XX
AC ABL58698;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human kinase encoding cDNA.
XX
KW Human; kinase protein; calcium/calmodulin-dependant protein kinase; eye;
KW retinoblastoma; brain; kinase modulator; gene; SNP;
KW single nucleotide polymorphism; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..172
FT CDS /*tag= a
FT /*tag= b
FT /*tag= c
FT /product= "human kinase"
FT variation replace(1296, G)
FT 3'UTR /*tag= c
FT /standard_name= "single nucleotide polymorphism"
FT 1691..2190
FT /*tag= d
WO200224920-A2.
28-MAR-2002.
19-SEP-2001; 2001WO-US029161.
19-SEP-2000; 2000US-0234493P.
13-NOV-2000; 2000US-0247031P.
06-DEC-2000; 2000US-00729995.
(PEKE) PE CORP NY.
Beasley EM, Wei M, Bonazzi VR, Sanders R, Di Francesco V;
WPI: 2002-404955/43.
P-PSDB; AB883054.
Novel peptide designated as human kinase useful as target for diagnosing

a disease or predisposition to the disease mediated by the peptide.

Claim 4b; Fig 1; 89pp; English.

The invention relates to an isolated peptide designated human kinase (HK), that has homology to members of the calcium/calmodulin-dependent protein kinase kinase subfamily. The mechanism of action of the protein of the invention is that of a kinase modulator. The human kinase of the invention is useful for creating a pharmaceutical composition for treating a disease or condition mediated by the human kinase. HK is also useful to provide a target for diagnosing a disease or predisposition to disease mediated by HK, and is also useful in pharmacogenic analysis. HK is useful for treating a disorder characterised by absence of an inappropriate or unwanted expression of HK, also as an immunogen to raise antibodies by administering HK to a mammalian organism e.g. rat, rabbit or mouse. Nucleic acids of the invention are useful as hybridisation probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells expressing a native HK are useful for assaying compounds that stimulate or inhibit HK function. Nucleic acids of the invention are also useful for producing transgenic animals. Experimental data indicates that kinase proteins of the present invention are expressed in humans in the eye (retinoblastomas) and brain. The current sequence represents cDNA encoding the human kinase of the invention

Sequence 2190 BP; 457 A; 655 C; 708 G; 370 T; 0 U; 0 Other;

Query Match 100.0%; Score 1515; DB 6; Length 2190;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGGGGGGTCCAGCTGTCTGTCGACGATCCTCGGGCAGAGCTGGTAGACGGGTG 60
DB |||
QY 173 ATGAGGGGGGTCCAGCTGTCTGTCGACGATCCTCGGGCAGAGCTGGTAGACGGGTG 232
DB |||
QY 61 GCAGCCATCATGATGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAC 120
DB |||
QY 233 GCAGCCATCATGATGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAC 292
DB |||
QY 121 GGTGTGACCCCCACACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB |||
QY 293 GGTGTGACCCCCACACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 352
DB |||
QY 181 CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 240
DB |||
QY 353 CTGCTCCAGCCCGGCTTAGCCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 412
DB |||
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGGCTGCCAGCCATCTCC 300
DB |||
QY 413 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGGCTGCCAGCCATCTCC 472
DB |||
QY 301 CCCGGGCTTGGCGAGGCGCCACCATCTGAGTCCCAACAGTGGCCATCTCAGATGCGAG 360
DB |||
QY 473 CCCGGGCTTGGCGAGGCGCCACCATCTGAGTCCCAACAGTGGCCATCTCAGATGCGAG 532
DB |||
QY 361 GACTGCTGAGCTGAGTCAACAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 420
DB |||
QY 533 GACTGCTGAGCTGAGTCAACAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 592
DB |||
QY 421 GTGCTGAGCTGAGTCAACAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 480
DB |||
QY 593 GTGCTGAGCTGAGTCAACAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 652
DB |||
QY 481 AAAAGAAGTCTACTGAAGCAGTATGCTTTTCCAGTCCGCTCCCGGAGGGTCCCGAG 540
DB |||
QY 653 AAAAGAAGTCTACTGAAGCAGTATGCTTTTCCAGTCCGCTCCCGGAGGGTCCCGAG 712
DB |||
QY 541 GTGCTCCAGGAGGACCAAGCAGTCTGCTGCCCTTGGAGCGGGTGTACAGAGAGATT 600
DB |||
QY 713 GCTGCCAGGAGGACCAAGCAGTCTGCTGCCCTTGGAGCGGGTGTACAGAGAGATT 772
DB |||
QY 601 GCATCTCAAGAGCTGAGCAGCAGTGAATGGTCAAACTGATCAGGTCTCTGGATGAC 660
DB |||
QY 773 GCATCTCAAGAGCTGAGCAGCAGTGAATGGTCAAACTGATCAGGTCTCTGGATGAC 832
DB |||
```

RESULT 2

AAD30566

ID AAD30566 standard; cDNA; 1937 BP.

XX

AC AAD30566;

XX

DT 21-MAY-2002 (first entry)

DE

XX Human kinase polypeptide (PKIN-19) cDNA.

XX

KW Human; kinase polypeptide; PKIN-19; gene therapy; Addison's disease;

KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; Bursitis; atherosclerosis;
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiant; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme; ss.

XX Homo sapiens.

OS
XX
FH Key Location/Qualifiers
CDS 65..1582
FT /*tag= a
FT /product= "Human PKIN-19"

PN WO200208399-A2.

XX 31-JAN-2002.

XX 20-JUL-2001; 2001WO-US023092.

XX 21-JUL-2000; 2000US-0220038P.

PR 28-JUL-2000; 2000US-0222112P.

PR 04-AUG-2000; 2000US-0222831P.

PR 11-AUG-2000; 2000US-0224729P.

XX (INCY-) INCYTE GENOMICS INC.

XX (THOR/) THORNTON M.

XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AUA, Walia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;
PI Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;
PI Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;

XX WPI; 2002-0206083/26.

DR P-PSDB; AAE19161.

XX New human kinase polypeptide, useful in diagnosis, prevention and
PT treatment of cancer, immune disorder, growth and developmental disorder,
PT cardiovascular disorder and lipid disorder.

XX Claim 5; Page 193-194; 196pp; English.

XX The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
CC present sequence is human PKIN-19 cDNA

XX Sequence 1937 BP; 425 A; 567 C; 603 G; 342 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 99.7%; Score 1510.2; DB 6; Length 1937;

XX Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTGTCTCCAGGATCCCTCGGCGAGAGCTGTGTAGAACGGGTG 60
DB 65 ATGGAGGGGGTCCAGCTGTCTGTCTCCAGGATCCCTCGGCGAGAGCTGTGTAGAACGGGTG 124
QY 61 GGAGCCATCGATGTGACTCACTTTGGAGGAGGAGATGGTGGCCAGAGCTTACTAGAAAC 120
DB 125 GGAGCCATCGATGTGACTCACTTTGGAGGAGGAGATGGTGGCCAGAGCTTACTAGAAAC 184
QY 121 GGTGTGACCCCCCACCAGGGCCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 180
DB 185 GGTGTGACCCCCCACCAGGGCCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
DB 245 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 304
QY 241 GGAGCTATCTGGAGCGCAGGCTGGGCCCTTATTCACGGGGCTTCGACGCCACATCTCC 300
DB 305 GGAGCTATCTGGAGCGCAGGCTGGGCCCTTATTCACGGGGCTTCGACGCCACATCTCC 364
QY 301 CCCGGGCTGGCGGAGGCCACCATCGAGTCCCACCACGCTGGCCATCTCAGATGAGAG 360
DB 365 CCCGGGCTGGCGGAGGCCACCATCGAGTCCCACCACGCTGGCCATCTCAGATGAGAG 424
QY 361 GACTGCGTGCAGCTGAACAGTACAAGCTGCGAGTGAAGTTGGCAAGGGTGCCTACGGT 420
DB 425 GACTGCGTGCAGCTGAACAGTACAAGCTGCGAGTGAAGTTGGCAAGGGTGCCTACGGT 484
QY 421 GTGGTGAGGCTGGCTTACAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
DB 485 GTGGTGAGGCTGGCTTACAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 544
QY 481 AAAAAAGATTACTGAAGCAGTATGGCTTTTCCAGCTGCGCCCTCCCCCGAGAGGGTCCCAG 540
DB 545 AAAAAAGATTACTGAAGCAGTATGGCTTTTCCAGCTGCGCCCTCCCCCGAGAGGGTCCCAG 604
QY 541 GCTGCCAGGGAGGACCAAGCAGAGCTGTGCTCCCTCGAGGGGTGTACCAGGAGATT 600
DB 605 GCTGCCAGGGAGGACCAAGCAGAGCTGTGCTCCCTCGAGGGGTGTACCAGGAGATT 664
QY 601 GCATCTCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGANTCAGAGTCTCTGGATGAC 660
DB 665 GCATCTCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGANTCAGAGTCTCTGGATGAC 724
QY 661 CCAGCTGAGGACAACTCTATTGTGTGTGACTCTCTGAGAAAGGGCCCGTCATGGAA 720
DB 725 CCAGCTGAGGACAACTCTATTGTGTGTGACTCTCTGAGAAAGGGCCCGTCATGGAA 784
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTGGGGACGTCATC 780
DB 785 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTGGGGACGTCATC 844
QY 781 CTGGGCTCGAGTACTTGCACTGCGAGAGATGCTCCAGGAGACATCAAGCCATCCAAC 840
DB 845 CTGGGCTCGAGTACTTGCACTGCGAGAGATGCTCCAGGAGACATCAAGCCATCCAAC 904
QY 841 CTGCTCTCGGGGATCATGGGCACGTTGAAGATGCCGACTTTGGCGTTCAGCAACCAAGTTT 900
DB 905 CTGCTCTCGGGGATCATGGGCACGTTGAAGATGCCGACTTTGGCGTTCAGCAACCAAGTTT 964
QY 901 GAGGGAAACAGCGCTCAGCTGTCCAGCACGGGGGAAACCCAGCAATTCATGGCCCCCGAG 960
DB 965 GAGGGAAACAGCGCTCAGCTGTCCAGCACGGGGGAAACCCAGCAATTCATGGCCCCCGAG 1024
QY 961 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCCCTTGAATGTATGGGCCACTGGC 1020
DB 1025 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCCCTTGAATGTATGGGCCACTGGC 1084
QY 1021 GTACGTTGCTGCTTTGCTTATGGGAAGTCCCATTCATCGAGATTTTCATCTCGGCC 1080
DB 1085 GTACGTTGCTGCTTTGCTTATGGGAAGTCCCGCTTCATCGAGATTTTCATCTCGGCC 1144

QY 1081 CTCACAGAGATCAAGATCAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1140
DB 1145 CTCACAGAGATCAAGATCAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1204
QY 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAAATTCGGGTG 1200
DB 1205 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAAATTCGGGTG 1264
QY 1201 CCAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGGAGGACCCCTTCTCTTCGGAG 1260
DB 1265 CCAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGGAGGACCCCTTCTCTTCGGAG 1324
QY 1261 GAGGAGCACTCAGCGTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB 1325 GAGGAGCACTCAGCGTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1384
QY 1321 ATCCCCAGCTGGAACCAAGCTGATCTGGTGAAGTCCATGCTGAGGAAGCTTCCTTTGGG 1380
DB 1385 ATCCCCAGCTGGAACCAAGCTGATCTGGTGAAGTCCATGCTGAGGAAGCTTCCTTTGGG 1444
QY 1381 AACCGTTTGAAGCCCAAGGACGAGGAGGAAGCGATCATGTCTGCTCCAGGAACCTA 1440
DB 1445 AACCGTTTGAAGCCCAAGGACGAGGAGGAAGCGATCATGTCTGCTCCAGGAACCTA 1504
QY 1441 CTGGTGAAGAGGCTTTGGTGAAGGGGGCAGAGCCAGAGCTCCCGGGCTCCAGGA 1500
DB 1505 CTGGTGAAGAGGCTTTGGTGAAGGGGGCAGAGCCAGAGCTCCCGGGCTCCAGGA 1564
QY 1501 GACGAGGCTGCATCC 1515
DB 1565 GACGAGGCTGCATCC 1579
RESULT 3
ABZ11551
ID ABZ11551 standard; cDNA; 2018 BP.
XX AC ABZ11551;
XX DT 20-JAN-2003 (first entry)
XX DE Human polynucleotide SEQ ID NO 433.
XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX OS Homo sapiens.
XX PN WO200270539-A2.
XX PD 12-SEP-2002.
XX PF 05-MAR-2002; 2002WO-US005095.
XX PR 05-MAR-2001; 2001US-00799451.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao Q, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI: 2002-759812/82.
XX DR P-PSDB; ABP69334.
XX PT New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative.

PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
or coagulation disorders.
XX Claim 1; SEQ ID NO 433; 1012pp + Sequence Listing; English.
XX The invention relates to an isolated polynucleotide (I) comprising a
nucleotide sequence selected from any of 948 sequences (ABZ11119-
ABZ12066) or their mature protein coding portion, active domain coding
protein or complementary sequences. The polynucleotides are useful for
identifying expressed genes or for physical mapping of human genome. The
encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight
markers, as a food supplement, for generating antibodies, in medical
imaging, screening and diagnostic assays and for treating cell
proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
platelet or coagulation disorders, wound, burns, incision, ulcers, liver
or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
arthritis, etc. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2018 BP; 453 A; 600 C; 603 G; 362 T; 0 U; 0 Other;
Query Match 99.5%; Score 1507; DB 6; Length 2018;
Best Local Similarity 99.7%; Fred. No. 0;
Matches 1510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGAGGGGGGTCCAGCTGTCTGTCGCGAGATCTTCGGGCGAGACTGGTAGAAGGGTG 60
DB 232 ATGAGGGGGGTCCAGCTGTCTGTCGCGAGATCTTCGGGCGAGACTGGTAGAAGGGTG 291
QY 61 GCAGCCATCGATGTGACTCACTTGAGAGGAGAGATGTGGCCGAGACCTACTAGAAAC 120
DB 292 GCAGCCATCGATGTGACTCACTTGAGAGGAGAGATGTGGCCGAGACCTACTAGAAAC 351
QY 121 GGTGTGACCCCGCCACACCGGCGAGAGCTGCTCTGTGATCCCTGGGAGTACTTCAAGA 180
DB 352 GGTGTGACCCCGCCACACCGGCGAGAGCTGCTCTGTGATCCCTGGGAGTACTTCAAGA 411
QY 181 CTGCTCCCGAGCCCGCTAGGCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCAGCA 240
DB 412 CTGCTCCCGAGCCCGCTAGGCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCAGCA 471
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCAGGGGCTTCGCCACATCTCC 300
DB 472 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCAGGGGCTTCGCCACATCTCC 531
QY 301 CCCGGGCTCGGGAGGCGCCACCATCGAGTCCACACCGTGGCCATCTCAGATCAGAG 360
DB 532 CCCGGGCTCGGGAGGCGCCACCATCGAGTCCACACCGTGGCCATCTCAGATCAGAG 591
QY 361 GACTGCGTGAGCTGAACCACTAGTACAGCTGAGAGTGGCAAGGGTGCCTACGGT 420
DB 592 GACTGCGTGAGCTGAACCACTAGTACAGCTGAGAGTGGCAAGGGTGCCTACGGT 651
QY 421 GTGTGAGGCTGGCCCTACAAAGAAAGTGAAGACAGACTATGCAATGAAGTCTTTCC 480
DB 652 GTGTGAGGCTGGCCCTACAAAGAAAGTGAAGACAGACTATGCAATGAAGTCTTTCC 711
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCCTCCCGGAGGGTCCCGAG 540
DB 712 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCCTCCCGGAGGGTCCCGAG 771
QY 541 GCTGCCAGGAGGAGCAGCCAGAGCTGTGCTGCCCTGGAGCGGGTGTACAGAGAGATT 600
DB 772 GCTGCCAGGAGGAGCAGCCAGAGCTGTGCTGCCCTGGAGCGGGTGTACAGAGAGATT 831
QY 601 GCCATCTGAAGAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGATGAC 660
DB 832 GCCATCTGAAGAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGATGAC 891
QY 661 CCAGCTGAGGACAACTCTATTGTTGTTGACCTCTCTGAGAAAGGGGCCGCTCATGGAA 720

Db 892 CCAGCTGAGGACAACTCTATTGTGTGTTGACCTCTGTGAAAAGGGCCCGTCATGGAA 951
Qy 721 GTGCCCTGTGCAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTCGCGGAGCTGATC 780
Db 952 GTGCCCTGTGCAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTCGCGGAGCTGATC 1011
Qy 781 CTGGCCCTCGAGTACTTGTCACTGTCAGAAAGATCGTCCACAGGACATCAAGCCATCCAAAC 840
Db 1012 CTGGCCCTCGAGTACTTGTCACTGTCAGAAAGATCGTCCACAGGACATCAAGCCATCCAAAC 1071
Qy 841 CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGGACTTTGGGCTCAGCAACAGATTT 900
Db 1072 CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGGACTTTGGGCTCAGCAACAGATTT 1131
Qy 901 GAGGGGAACGAGCTGAGCTGTCAGACGCGCGGGAACCCAGCATTCATGCGCCCGCGAG 960
Db 1132 GAGGGGAACGAGCTGAGCTGTCAGACGCGCGGGAACCCAGCATTCATGCGCCCGCGAG 1191
Qy 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1020
Db 1192 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1251
Qy 1021 GTACGTTGTACTGCTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCATCTGGCC 1080
Db 1252 GTACGTTGTACTGCTTGTCTATGGGAAGTGCCCGTTTCATCGACGATTTTCATCTGGCC 1311
Qy 1081 CTCACAGGAAGATCAAGAATGAGCCGTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1312 CTCATAGGAAGATCAAGAATGAGCCGTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1371
Qy 1141 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGAAATCCCGAGAGAGAAATTTGGGTTG 1200
Db 1372 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGAAATCCCGAGAGAGAAATTTGGGTTG 1431
Qy 1201 CCAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGGAGAGCCCTTCCTTCGGAG 1260
Db 1432 CCAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGGAGAGCCCTTCCTTCGGAG 1491
Qy 1261 GAGGAGCACTGACGCTGGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1492 GAGGAGCACTGACGCTGGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1551
Qy 1321 ATCCCGAGCTGACCAAGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTTCCTTTGGG 1380
Db 1552 ATCCCGAGCTGACCAAGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTTCCTTTGGG 1611
Qy 1381 AACCCGTTGAGCCCGGAGGACGAGGAGGAGCGATCCATGCTGCTCCAGGAACCTA 1440
Db 1612 AACCCGTTGAGCCCGGAGGACGAGGAGGAGCGATCCATGCTGCTCCAGGAACCTA 1671
Qy 1441 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGAGAA 1500
Db 1672 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGAGAA 1731
Qy 1501 GAGGAGCTGCATCC 1515
Db 1732 GACGAGGCTGCATCC 1746

RESULT 4
ABX97035

ID ABX97035 standard; cDNA; 1611 BP.

XX AC ABX97035;

XX DT 20-MAY-2003 (first entry)

XX DE Human NOV15b cDNA.

KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.

XX OS Homo sapiens.
XX PN WO200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-027791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294489P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332694P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX

(CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
XX Zernhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JB;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR P-PSDB; ABU65068.
XX

NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.

XX Claim 13; Page 139; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
CC cytotstatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX

XX Sequence 1611 BP; 366 A; 461 C; 497 G; 287 T; 0 U; 0 Other;

Query Match 99.4%; Score 1505.4; DB 6; Length 1611;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAACCGGGTG 60
DB |||||
52 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAACCGGGTG 111
QY 61 GGAGCCATCGATGACTGACTGAGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 120
DB |||||
112 GGAGCCATCGATGACTGACTGAGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 171
QY 121 GGTGTGGACCCGCCACCGGSCAGAGCTGCTGTGATCCCTGGCAGTACTTCAAGA 180
DB |||||
172 GGTGTGGACCCGCCACCGGSCAGAGCTGCTGTGATCCCTGGCAGTACTTCAAGA 231
QY 181 CTGCTCCCAAGCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
DB |||||
232 CTGCTCCCAAGCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 291
QY 241 GGAAGTATCTGGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGGCCACATCTCC 300
DB |||||
292 GGAAGTATCTGGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGGCCACATCTCC 351
QY 301 CCCCAGGCTGGCGAGGCGCCACCATCGAGTCCACACGCTGGCCATCTCAGATGCAGAG 360
DB |||||
352 CCCCAGGCTGGCGAGGCGCCACCATCGAGTCCACACGCTGGCCATCTCAGATGCAGAG 411
QY 361 GACTCGTGCAGCTGAACAGGACAGCTGAGCTGCAGAGTGAGATTGGCAAGGGTGCTACGGT 420
DB |||||
412 GACTCGTGCAGCTGAACAGGACAGCTGAGCTGCAGAGTGAGATTGGCAAGGGTGCTACGGT 471
QY 421 GTGGTGAGGCTGGCTAGAACCAAGTGAAGCAGACACTNTGCAATGAAGTCTTTCC 480
DB |||||
472 GTGGTGAGGCTGGCTAGAACCAAGTGAAGCAGACACTATGCAATGAAGTCTTTCC 531
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTTCCACGTCGCGCTCCCGGAGAGGGTCCACAG 540
DB |||||
532 AAAAAGAGTTACTGAAGCAGTATGGCTTTTCCACGTCGCGCTCCCGGAGAGGGTCCACAG 591

RESULT 5
ABK49563
ID ABK49563 standard; DNA; 2711 BP.
XX

QY 541 GCTGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 600
DB |||||
592 GTGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 651
QY 601 GCCATCTGAAGAAGCTGGACCACTGTAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660
DB |||||
652 GCCATCTGAAGAAGCTGGACCACTGTAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 711
QY 661 CCAGCTGAGGACAACTCTATTGTTGTTGACCTCTCAGAAAGGGGCCGTCATGGAA 720
DB |||||
712 CCAGCTGAGGACAACTCTATTGTTGTTGACCTCTCAGAAAGGGGCCGTCATGGAA 771
QY 721 GTGCCCTGTGACAAAGCCCTCTCGAGAGGAGCAAGCTCGCTCTACCTGCGGAGCTCATC 780
DB |||||
772 GTGCCCTGTGACAAAGCTCTCTCGAGAGGAGCAAGCTCGCTCTACCTGCGGAGCTCATC 831
QY 781 CTGGGCTCGAGTACTTGCACTGCCAGAAAGATCGTCCACAGGAGACATCAAGCCATCCAAC 840
DB |||||
832 CTGGGCTCGAGTACTTGCACTGCCAGAAAGATCGTCCACAGGAGACATCAAGCCATCCAAC 891
QY 841 CTGCTCTGGGGATGATGGGCACGCTGAAGATCGCGACTTTGGCGTCAGCAACCAAGTTT 900
DB |||||
892 CTGCTCTGGGGATGATGGGCACGCTGAAGATCGCGACTTTGGCGTCAGCAACCAAGTTT 951
QY 901 GAGGGGAACGACGCTCAGCTGCCAGCAGCGGGGAACCCCAAGCATTCATGGCCCCCGAG 960
DB |||||
952 GAGGGGAACGACGCTCAGCTGCCAGCAGCGGGGAACCCCAAGCATTCATGGCCCCCGAG 1011
QY 961 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGATGAGGCACTGGC 1020
DB |||||
1012 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGATGAGGCACTGGC 1071
QY 1021 GTACGTTGCTACTGCTTCTGCTATGGAGTGCCTCATTCATGACGATTCATCTCTGGCC 1080
DB |||||
1072 GTACGTTGCTACTGCTTCTGCTATGGAGTGCCTGTTTATGACGATTCATCTCTGGCC 1131
QY 1081 CTCCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTGAGAGGCGAGAAATCAGCGAG 1140
DB |||||
1132 CTCCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTGAGGGGCCAGAAATCAGCGAG 1191
QY 1141 GAGCTCAGGAGCTGATCTGAAAGATGTTAGACAGAAATCCCGAGACGAGAAATGGGGTG 1200
DB |||||
1192 GAGCTCAGGAGCTGATCTGAAAGATGTTAGACAGAAATCCCGAGACGAGAAATGGGGTG 1251
QY 1201 CCAGACATCAAGTTCACACCTTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAG 1260
DB |||||
1252 CCAGACATCAAGTTCACACCTTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAG 1311
QY 1261 GAGGAGCACTGAGCGTGGTGGAGGTGACAGAGGGGGAGGTTTAAAGAACTCAGTCAGGCTC 1320
DB |||||
1312 GAGGAGCACTGAGCGTGGTGGAGGTGACAGAGGAGGAGGTTTAAAGAACTCAGTCAGGCTC 1371
QY 1321 ATCCCCAGCTGACACACGCTGATCTGTTGAAGTCCATGCTCAGGAGAGGCTTCTTTGGG 1380
DB |||||
1372 ATCCCCAGCTGACACACGCTGATCTGTTGAAGTCCATGCTCAGGAGAGGCTTCTTTGGG 1431
QY 1381 AAACCGTTTGGCCCGCAGGACGAGGAGGAGCGATCCATGCTCTCCAGAAACCTTA 1440
DB |||||
1432 AAACCGTTTGGCCCGCAGGACGAGGAGGAGGAGCGATCCATGCTCTCCAGAAACCTTA 1491
QY 1441 CTGGTGAAGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCCGGCGTCCAGGAA 1500
DB |||||
1492 CTGGTGAAGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCCGGCGTCCAGGAA 1551
QY 1501 GACGAGGCTGCATCC 1515
DB |||||
1552 GACGAGGCTGCATCC 1566

AC ABK49563;
XX 15-JUL-2002 (first entry)
XX
DE Human cDNA 16002 encoding a novel kinase.
XX
KW Human; ss; gene; 16002; kinase; cellular proliferative disorder; cancer;
KW carcinoma; tumour; adenocarcinoma; haematopoietic neoplastic disorder;
KW leukaemia; lymphoma; brain disorder; cerebral ischaemia; infection;
KW meningitis; brain abscess; acquire immunodeficiency syndrome; obesity;
KW AIDS-related myopathy; prion disease; Alzheimer's disease; diabetes;
KW Parkinson's disease; Huntington's disease; motor neurone disease;
KW metabolic disorder; anorexia nervosa; pain; inflammation; ischaemia;
KW irritable bowel syndrome; heart disorder; myocardial infarction;
KW blood vessel disorder; atherosclerosis; bone metabolism disorder;
KW osteoporosis; haematopoietic disorder; arthritis.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
FH Key 198..1883
FT CDS /*tag= a
FT /product= "Kinase 16002"
FT /note= "this coding sequence (minus the stop codon) is
FT specifically claimed in claim 1"
XX
XX WO200220800-A2.
XX
XX 14-MAR-2002.
XX
XX 03-AUG-2001; 2001WO-US024601.
XX
XX 01-SEP-2000; 2000US-0229299P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers RA, Silos-Santiago I;
XX WPI; 2002-351781/38.
XX P-PSDB; AAU79458.
XX
XX New human transferase protein for diagnosing and treating disorders e.g.
XX cancer, Alzheimer's disease, anorexia, diabetes and to identify
XX modulators for therapeutic use.
XX
XX Claim 1; Fig 16; 143pp; English.
XX
XX The invention relates to an isolated human kinase polypeptide encoded by
XX the DNAs designated 16658, 14223 and 16002 including fragments,
XX homologues and allelic variants. Also included are a host cell comprising
XX the DNA, an antibody which selectively binds to the novel kinase, a
XX method for producing the novel kinase comprising culturing the host cell
XX and recovering the protein, detecting the presence of the DNA in a sample
XX comprising contacting the sample with a compound which selectively
XX hybridises to the DNA and determining whether the compound has bound and
XX identifying compounds which bind to and/or modulate the protein
XX comprising contacting the protein with a test compound and determining
XX whether the compound has bound to and/or modulated the function of the
XX protein (the modulators may be a small molecule, a peptide, a
XX phosphopeptide, an antibody or a fragment of the full length protein).
XX The nucleic acids, proteins, identified modulators and antibodies are
XX useful in the diagnosis, monitoring and treatment of a wide range of
XX diseases and disorders (many examples of which are listed in the
XX specification) including cellular proliferative disorders (e.g. cancers
XX of the lung and breast, carcinomas, tumours, adenocarcinomas,
XX haematopoietic neoplastic disorders e.g. leukaemias and lymphomas), brain
XX and nerve tissue disorders (e.g. cerebral ischaemia, infections such as
XX meningitis, brain abscess, acquired immunodeficiency syndrome (AIDS) -
XX related myopathy, prion disease, Alzheimer's disease, Parkinson's
XX disease, Huntington's disease and motor neurone disease) metabolic
XX disorders (e.g. obesity, anorexia nervosa and diabetes) pain disorders
XX (e.g. associated with infection, inflammation ischaemia, irritable bowel
XX syndrome), heart disorders (e.g. myocardial infarction), blood vessel

CC disorders (e.g. atherosclerosis), disorders of bone metabolism (e.g.
CC osteoporosis) and haematopoietic disorders (e.g. arthritis). The present
CC sequence encodes novel human kinase 16002
XX
SQ Sequence 2711 BP; 592 A; 752 C; 848 G; 508 T; 0 U; 11 Other;

Query Match 98.4%; Score 1490.6; DB 6; Length 2711;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTCTGCCAGGATCTCGGCGAGCTGGTAGAACGGTG 60
DB 198 ATGGAGGGGGTCCAGCTGTCTCTGCCAGGATCTCGGCGAGCTGGTAGAACGGTG 257
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGAGTGGTGGCCAGAGCCTACTAGAAAC 120
DB 258 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGAGTGGTGGCCAGAGCCTACTAGAAAC 317
QY 121 GGTGTGAGACCCCCACACAGGGCCAGAGTGCCTCTGTGTATCTCTGCGCATCTTCAAGA 180
DB 318 GGTGTGAGACCCCCACACAGGGCCAGAGTGCCTCTGTGTATCTCTGCGCATCTTCAAGA 377
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
DB 378 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 437
QY 241 GGAAGCTATCTGGAGCGCAGGCTGGGCGCTTATGCCACGGGGCTGCCAGCCACATCTCC 300
DB 438 GGAAGCTATCTGGAGCGCAGGCTGGGCGCTTATGCCACGGGGCTGCCAGCCACATCTCC 497
QY 301 CCCCGGCTGGCGAGGCGCCACCATCGAGTCCACACAGCTGGCCATCTCAGATGAGAG 360
DB 498 CCCCGGCTGGCGAGGCGCCACCATCGAGTCCACACAGCTGGCCATCTCAGATGAGAG 557
QY 361 GACTGCGTGCAGCTGAACAGTACAGCTGCAGAGTGAATTCGCAAGGCTGCTAGCGT 420
DB 558 GACTGCGTGCAGCTGAACAGTACAGCTGCAGAGTGAATTCGCAAGGCTGCTAGCGT 617
QY 421 GTGCTGAGGCTGGCTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTC 480
DB 618 GTGCTGAGGCTGGCTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTC 677
QY 481 AAAAAGAGTTACTGAAGAGTATGGCTTTCCAGTCCGCTCCCGGAGAGGTTCCAG 540
DB 678 AAAAAGAGTTACTGAAGAGTATGGCTTTCCAGTCCGCTCCCGGAGAGGTTCCAG 737
QY 541 GCTGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGCGGTTGACAGAGATT 600
DB 738 GCTGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGCGGTTGACAGAGATT 797
QY 601 GCCATCTCTGAAGAAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 660
DB 798 GCCATCTCTGAAGAAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 857
QY 661 CAGCTGAGGAGCAACCTCTATTTGGTGTGGAGTCTCTGAGAAAGGGGCCGCTATGAGAA 720
DB 858 CAGCTGAGGAGCAACCTCTATTTGGTGTGGAGTCTCTGAGAAAGGGGCCGCTATGAGAA 917
QY 721 GTCCCTGTGACAAAGCCCTCTCGAGGAGCAGCTGCGCTTACCTGCGGGACGTATC 780
DB 918 GTCCCTGTGACAAAGCCCTCTCGAGGAGCAGCTGCGCTTACCTGCGGGACGTATC 977
QY 781 CTGGGCTCGAGTACTTGCATCTGCCAGAGATCGTCCAGGGAGCATCAAGCCATCCAAC 840
DB 978 CTGGGCTCGAGTACTTGCATCTGCCAGAGATCGTCCAGGGAGCATCAAGCCATCCAAC 1037
QY 841 CTGCTCTCGGGGATGATGGGCGAGTGAAGATCGCGAGCTTTGGGTCAGCAACAGTTT 900
DB 1038 CTGCTCTCGGGGATGATGGGCGAGTGAAGATCGCGAGCTTTGGGTCAGCAACAGTTT 1097
QY 901 GAGGGNACACCGCTCAGCTGTCCAGCAGCGGGGAGCCCGACAGCATTCATGGCCCCGAG 960
DB 1098 GAGGGNACACCGCTCAGCTGTCCAGCAGCGGGGAGCCCGACAGCATTCATGGCCCCGAG 1157

QY 961 GCATTCTTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGCCACTGGC 1020
DB 1158 GCATTCTTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGCCACTGGC 1217
QY 1021 GTCACGTTGCTACTGCTTTGCTATGGGAAGTGCCTTATCATCAGCATTTTCATCCTTGGCC 1080
DB 1218 GTCACGTTGCTACTGCTTTGCTATGGGAAGTGCCTTATCATCAGCATTTTCATCCTTGGCC 1277
QY 1081 CTTCCACAGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
DB 1278 CTTCCACAGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1337
QY 1141 GAGCTCAAGGACTGATCCTGAAGATGTTAGCAAGAAATCCCGAGAGCAATGGGGTG 1200
DB 1338 GAGCTCAAGGACTGATCCTGAAGATGTTAGCAAGAAATCCCGAGAGCAATGGGGTG 1397
QY 1201 CCAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGAGGAGCCCTTCTTCCGAG 1260
DB 1398 CCAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGAGGAGCCCTTCTTCCGAG 1457
QY 1261 GAGGACATGACGCTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB 1458 GAGGACATGACGCTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1517
QY 1321 ATCCCCAGCTGGACACAGCTGATCCTGGTGAAGTCCATCTGAGGAAGCGTTCCTTGGG 1380
DB 1518 ATCCCCAGCTGGACACAGCTGATCCTGGTGAAGTCCATCTGAGGAAGCGTTCCTTGGG 1577
QY 1381 AACCCGTTTGACCCCGACGAGGAGGAGCGATCCATCTGTCTCCAGGAAACCTA 1440
DB 1578 AACCCGTTTGACCCCGACGAGGAGGAGCGATCCATCTGTCTCCAGGAAACCTA 1637
QY 1441 CTGGTGAAGAGGCTTTGGTGAAGGGGCGAGAGCCAGAGCTCCCGGCGTCCAGAA 1500
DB 1638 CTGGTGAAGAGGCTTTGGTGAAGGGGCGAGAGCCAGAGCTCCCGGCGTCCAGGCT 1697
QY 1501 GACGA 1505
DB 1698 TACCA 1702
RESULT 6
AAS06710
ID AAS06710 standard; cDNA; 1542 BP.
AC AAS06710;
XX
DT 12-SEP-2001 (first entry)
DE Polynucleotide sequence encoding human protein kinase #10.
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX WO200138503-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032085.
XX
XX 24-NOV-1999; 99US-0167482P.
XX
XX (SUGB-) SUGEN INC.
XX
XX Flowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX Flanagan P, Clary D;
XX
XX WPI; 2001-343950/36.
DB

DR P-PSDB; AAU03510.
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
XX Example 1; Fig 1; 433pp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 1542 BP; 349 A; 428 C; 483 G; 282 T; 0 U; 0 Other;
Query Match 96.3%; Score 1458.8; DB 4; Length 1542;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 2; Indels 30; Gaps 2;
QY 1 ATGAGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCGGGCAGAGCTGTAGAACGGGTG 60
DB 1 ATGAGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCGGGCAGAGCTGTAGAACGGGTG 60
QY 61 GCAGCCATCGATGTGACTCAGTCTGAGGAGGAGGAGTGTGGCCAGAGCTACTAGAAC 120
DB 61 GCAGCCATCGATGTGACTCAGTCTGAGGAGGAGGAGTGTGGCCAGAGCTACTAGAAC 120
QY 121 GGTGTGGACCCCCACCCAGCGGCTGTCTGTGATCTCTGGCAGTACTTCAAGA 180
DB 121 GGTGTGGACCCCCACCCAGCGGCGAGAGCTGCTCTGTGATCTCTGGCAGTACTTCAAGA 180
QY 181 CTGCTCCAGACCCGCGCTAGCTCTCAGCCAGGAAGTCTTCCCTACAGAGGGCGCAGCA 240
DB 181 CTGCTCCAGACCCGCGCTAGCTCTCAGCCAGGAAGTCTTCCCTACAGAGGGCGCAGCA 240
QY 241 GGAAGCTATCTGGAGCGCAGCTGGGCTTATGCCAGGGGCTGCCAGCCACATCTCC 300
DB 241 GGAAGCTATCTGGAGCGCAGCTGGGCTTATGCCAGGGGCTGCCAGCCACATCTCC 300
QY 301 CCCCGGGCTCGCGGAGGCCCAACATCGAGTCCACACCGTGGCCATCTCAGATCAGAG 360
DB 301 CCCCGGGCTCGCGGAGGCCCAACATCGAGTCCACACCGTGGCCATCTCAGATCAGAG 360
QY 361 GACTGCGTGCAGCTGAACCACTACAGCTGCAGAGTGAATGGCA----- 406
DB 361 GACTGCGTGCAGCTGAACCACTACAGCTGCAGAGTGAATGGCA----- 420
QY 407 -----AGGGTGCCTACGGTGTGTGAGGCTGGCCTACACGAAGTGAAGAC 453
DB 421 GATGCCTATCTGCAGGGTGCTTACGGTGTGTGAGGCTGGCCTACACGAAGTGAAGAC 480
QY 454 AGACACTATGCAATGAAAGTCTTTCAAAAAAGTGTACTGAAGCAGTATGGCTTTCCA 513
DB 481 AGACACTATGCAATGAAAGTCTTTCAAAAAAGTGTACTGAAGCAGTATGGCTTTCCA 540
QY 514 CGTGGCCTCCCGGAGAGGGTCCAGGCTCCCGGAGGAGGAGCCAGCCAGCAGTGTG 573
DB 541 CGTGGCCTCCCGGAGAGGGTCCAGGCTCCCGGAGGAGGAGCCAGCCAGCAGTGTG 600
QY 574 CCCCTGGAGCGGGTGTATACAGGAGATTTGCCATCTCTGAAGAGAGCTGGACCACTGAATGTG 633

Db 65 ATGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGCTAGAACGGGTG 124
QY 61 GCAGCCATGATGTGACTCACCTTGGAGGAGGAGATGGTGGCCACAGACCTACTAGAAC 120
Db 125 GCAGCCATGATGTGACTCACCTTGGAGGAGGAGATGGTGGCCACAGACCTACTAGAAC 184
QY 121 GGTGTGGACCCCCACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGGACCCCCACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGTCCCGACCGGGCTAGCCTCTCAGCCAGGAAGCTTTTCCTACAGGAGGGCCAGCA 240
Db 245 CTGTCCCGACCGGGCTAGCCTCTCAGCCAGGAAGCTTTTCCTACAGGAGGGCCAGCA 304
QY 241 GGAAGCTATCTGAGGCGCAGCTGGCCTTATGCCACGGGGCCTGCCAGCCACATCTCC 300
Db 305 GGAAGCTATCTGAGGCGCAGCTGGCCTTATGCCACGGGGCCTGCCAGCCACATCTCC 364
QY 301 CCCCAGGCTGCGGAGGGCCACCATCGAGTCCCACACGTCGTGGCCATCTCAGATGCAGAG 360
Db 365 CCCCAGGCTGCGGAGGGCCACCATCGAGTCCCACACGTCGTGGCCATCTCAGATGCAGAG 424
QY 361 GACTGGTGACGTGAACAGATCAAGCTGCGAGTGAGATTTGGCA----- 406
Db 425 GACTGGTGACGTGAACAGATCAAGCTGCGAGTGAGATTTGGCAAGGTGGGCTGACT 484
QY 407 -----AGGTCCTACGCTGCTGGTGAGGCTGGCCTACAACGAAAGTGAAGAC 453
Db 485 GATGCTATCTGAGGGTGCCTTACCGTGTGGTGGCTGGCCTACAACGAAAGTGAAGAC 544
QY 454 AGACACTATGCAATGAAAGTCTTTCCAAAGAAAGTTACTGAAGCAGTATGGCTTTCCA 513
Db 545 AGACACTATGCAATGAAAGTCTTTCCAAAGAAAGTTACTGAAGCAGTATGGCTTTCCA 604
QY 514 GGTGCGCTCCCGCAGAGGGTCCAGGCTGCCACGGAGGACCAAGCAAGCAGTGTGTG 573
Db 605 GGTGCGCTCCCGCAGAGGGTCCAGGCTGCCACGGAGGACCAAGCAAGCAGTGTGTG 664
QY 574 CCGCTGGAGCGGTGTACAGAGAGTTGCATCTCTGAAGAGAGCTGGACCAAGTGAATGTG 633
Db 665 CCGCTGGAGCGGTGTACAGAGAGTTGCCATCTCTGAAGAGAGCTGGACCAAGTGAATGTG 724
QY 634 GTCAAACCTGATCAGAGTCTGTGATGACCCAGCTGAGGACCAACTCTATTTGGTTTGAC 693
Db 725 GTCAAACCTGATCAGAGTCTGTGATGACCCAGCTGAGGACCAACTCTATTTGG---TTGAC 781
QY 694 CTCCTGAGAAAGGGCCCGTCAATGGAAGTGCCTGTGACAGCCCTTCTCGGAGAGCAA 753
Db 782 CTCCTGAGAAAGGGCCCGTCAATGGAAGTGCCTGTGACAGCCCTTCTCGGAGAGCAA 841
QY 754 GCTCGCCTCTACCTCGGGAGCTCATCTGGGCTTCGAGTACTTGCATGCCAGAGATC 813
Db 842 GCTCGCCTCTACCTCGGGAGCTCATCTGGGCTTCGAGTACTTGCATGCCAGAGATC 901
QY 814 GTCCACAGGACATCAAGCCATCCAACTGCTCTCGGGGATGATGGGCACGTTGAAGATC 873
Db 902 GTCCACAGGACATCAAGCCATCCAACTGCTCTCGGGGATGATGGGCACGTTGAAGATC 961
QY 874 GCGGACTTTGGCTCAGCAACAGTTTGGGGGACGACGCTCAGCTGTCAGACAGCGC 933
Db 962 GCGGACTTTGGCTCAGCAACAGTTTGGGGGACGACGCTCAGCTGTCAGACAGCGC 1021
QY 934 GGAACCCCGACATTCATGCCCCCGAGGCCATTTCTGATTCGGCCAGAGCTTCAGTGGG 993
Db 1022 GGAACCCCGACATTCATGCCCCCGAGGCCATTTCTGATTCGGCCAGAGCTTCAGTGGG 1081
QY 994 AAGGCTTTGGATGATGGGCCACTGGGCTCAGTTGTACTGCTTTGTCTATGGGAAGTC 1053
Db 1082 AAGGCTTTGGATGATGGGCCACTGGGCTCAGTTGTACTGCTTTGTCTATGGGAAGTC 1141
QY 1054 CCATTTCGACGATTTTCATCTGGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTG 1113
Db 1142 CCGTTTCGACGATTTTCATCTGGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTG 1201

QY 1114 TTTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC 1173
Db 1202 TTTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC 1261
QY 1174 AAGAATCCCGACGAGGATTCGGGTGCCAGACATCAAGTTTCACCCCTTGGGTGACCAAG 1233
Db 1262 AAGAATCCCGACGAGGATTCGGGTGCCAGACATCAAGTTTCACCCCTTGGGTGACCAAG 1321
QY 1234 AACGGGAGGAGCCCTTCTTCGGAGGAGGACACTGTCAGGCTGTGTGAGGTGACAGAG 1293
Db 1322 AACGGGAGGAGCCCTTCTTCGGAGGAGGACACTGTCAGGCTGTGTGAGGTGACAGAG 1381
QY 1294 GGGGAGGTTAAGAACTCAGTCAAGGCTCATCCCGAGTGGACCAAGTTCCTGTGTGAAG 1353
Db 1382 GAGGAGGTTAAGAACTCAGTCAAGGCTCATCCCGAGTGGACCAAGTTCCTGTGTGAAG 1441
QY 1354 TCCATGCTGAGGAAGCTTCTTTGGGAACCCGTTTGGAGCCCGAGCGAGGGAAGAG 1413
Db 1442 TCCATGCTGAGGAAGCTTCTTTGGGAACCCGTTTGGAGCCCGAGCGAGGGAAGAG 1501
QY 1414 CGATCCATGTCGTCTCCAGGAAACCTACTGTGTAAAGAGGTTTGGTGAAGGGGCAAG 1473
Db 1502 CGATCCATGTCGTCTCCAGGAAACCTACTGTGTAAAGAGGTTTGGTGAAGGGGCAAG 1561
QY 1474 AGCCGAGAGCTCCCGGGCTCCAGGAACGAGGCTGCATCC 1515
Db 1562 AGCCGAGAGCTCCCGGGCTCCAGGAAGCGAGGCTGCATCC 1603
RESULT 8
ABX97034
ID ABX97034 standard; cDNA; 1547 BP.
XX
AC ABX97034;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV15a cDNA.
XX
KW NOX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279959P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
XX (CURA-) CURAGEN CORP.
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
XX Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
XX Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
XX Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
XX Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
XX P-FSDB; ABU65067.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
XX a disorder associated with aberrant NOVX expression or activity e.g.,
XX cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
XX asthma.
XX
XX Claim 13; Page 138; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
XX cytotostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
XX activity. Pharmaceutical compositions comprising the NOVX proteins or
XX nucleic acid molecules or NOVX antibodies are useful for preventing or
XX treating a disorder associated with aberrant NOVX expression or activity
XX e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
XX asthma. The products of the invention can be used for gene therapy or in
XX a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
XX the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
XX ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
XX ABU65041-ABU65218

XX
SQ Sequence 1547 BP; 352 A; 441 C; 474 G; 280 T; 0 U; 0 Other;
Query Match 96.1%; Score 1455.8; DB 6; Length 1547;
Best Local Similarity 98.5%; Pred. No. 0; Mismatches 17; Indels 6; Gaps 2;
Matches 1492; Conservative 0;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCGAGAGTGTGTAGAACGGGTG 60
DB 20 ATGGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCGAGAGTGTGTAGAACGGGTG 79
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCTTACTAGAAC 120
DB 80 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCTTACTAGAAC 139
QY 121 GGTGTGACCCCCCACCAGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTCAAGA 180
DB 140 GGTGTGACCCCCCACCAGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTCAAGA 199
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 240
DB 200 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 259
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCAAGGGGCTGCGACCAATCTCC 300
DB 260 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCAAGGGGCTGCGACCAATCTCC 319
QY 301 CCCCAGGCTGCGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGAGAG 360
DB 320 CCCCAGGCTGCGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGAGAG 379
QY 361 GACTGCTGCAGCTGAACACAGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGGCTACGGT 420
DB 380 GACTGCTGCAGCTGAACACAGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGGCTACGGT 439
QY 421 GTGCTGAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAAGTCTTTTC 480
DB 440 GTGCTGAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAAGTCTTTTC 499
QY 481 AAAAGAAGTTACTGAAGCAGTATGCTTTTCCAGCTGCGCTCCCGGAGAGGTTCCAG 540
DB 500 AAAAGAAGTTACTGAAGCAGTATGCTTTTCCAGCTGCGCTCCCGGAGAGGTTCCAG 559
QY 541 GCTGCCAGGAGGAGGAGCCAGCAAGAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT 600
DB 560 GCTGCCAGGAGGAGGAGCCAGCAAGAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT 619
QY 601 GCATCTGTGAAGAGCTGGACACAGTGAATGTGTGTAAGTGAAGTGTGATGAGTGTGATGAC 660
DB 620 GCATCTGTGAAGAGCTGGACACAGTGAATGTGTGTAAGTGAAGTGTGATGAGTGTGATGAC 679
QY 661 CCAGCTGAGGACAACTCTATTGTTGTTGACCTCTCGAAGAGGGGCGGCTCATGGAA 720
DB 680 CCAGCTGAGGACAACTCTATTGTTGTTGACCTCTCGAAGAGGGGCGGCTCATGGAA 736
QY 721 GTCCCTCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTGCGCTCTACCTGCGGAGCGTCATC 780
DB 737 GTCCCTCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTGCGCTCTACCTGCGGAGCGTCATC 796
QY 781 CTGGGCTCTGAGTACTTGCATCTGCCAGAGAGATGCTCCACAGGAGACATCAAGCCATCAAC 840
DB 797 CTGGGCTCTGAGTACTTGCATCTGCCAGAGAGATGCTCCACAGGAGACATCAAGCCATCAAC 856
QY 841 CTGCTCTGGGGATGATGGGACAGTGAAGATCGCGACTTTTGGCGTACGAACACAGTTT 900
DB 857 CTGCTCTGGGGATGATGGGACAGTGAAGATCGCGACTTTTGGCGTACGAACACAGTTT 916
QY 901 GAGGGAAACGACGCTCAGCTGTCCAGCAGCGGCGGAAACCCAGCAATTCATGGCCCCCGAG 960
DB 917 GAGGGAAACGACGCTCAGCTGTCCAGCAGCGGCGGAAACCCAGCAATTCATGGCCCCCGAG 976
QY 961 GCATTTCTGATTCCTCGGCGAGAGCTTCACTGAGGAGAGGCTTGTGATGTATGGGCACTGGC 1020

Db 977 GCCATTTCTGATTCCGGCCAGAGCTTCACTGGGAAG---TTGGATGATATGGGCCACTGGC 1033
QY 1021 GTCAAGTTCTACTGCTTCTCTATGGGAAGTGCCTATTCATCGACCAATTTCTATCCTGGCC 1080
Db 1034 GTCAAGTTCTACTGCTTCTCTATGGGAAGTGCCTATTCATCGACCAATTTCTATCCTGGCC 1093
QY 1081 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1094 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1153
QY 1141 GAGCTCAAGGACTGATCTCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTGGGTTG 1200
Db 1154 GAGCTCAAGGACTGATCTCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTGGGTTG 1213
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCTTCGGAG 1260
Db 1214 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCTTCGGAG 1273
QY 1261 GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1274 GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1333
QY 1321 ATCCCCAGCTGACACAGGTGATCTCTGTGAAGTCCATCTCAGGAGAGGCTTCCTTTGGG 1380
Db 1334 ATCCCCAGCTGACACAGGTGATCTCTGTGAAGTCCATCTCAGGAGAGGCTTCCTTTGGG 1393
QY 1381 AACCCGTTTGAGCCCCAGGACGAGGAGGAAGAGCGATCCATCTCTGCCAGGAACCTA 1440
Db 1394 AACCCGTTTGAGCCCCAGGACGAGGAGGAAGAGCGATCCATCTCTGCCAGGAACCTA 1453
QY 1441 CTGGTGAAGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1500
Db 1454 CTGGTGAAGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1513
QY 1501 GACGAGGCTGCATCC 1515
Db 1514 GACGAGGCTGCATCC 1528

RESULT 9
ABX97036

ID ABX97036 standard; cDNA; 1725 BP.
XX
AC ABX97036;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV15c cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276766P.
PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.

(CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
XX P-PSDB; ABU65069.
DR
DR
XX
PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
XX asthma.
PS Claim 13; Page 139-140; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive

CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX
SQ Sequence 1725 BP; 394 A; 500 C; 523 G; 308 T; 0 U; 0 Other;
Query Match 91.2%; Score 1381.4; DB 6; Length 1725;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 6; Indels 114; Gaps 1;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTCCAGAGTCTCTGGGCGAGAGTGTGTAGAACGGGTG 60
DB 52 ATGGAGGGGGTCCAGCTGTCTGCTCCAGAGTCTCTGGGCGAGAGTGTGTAGAACGGGTG 111
QY 61 GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 120
DB 112 GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 171
QY 121 GGTGTGAGCCCCCACCACGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAAG 180
DB 172 GGTGTGAGCCCCCACCACGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAAG 231
QY 181 CTGCTCCAGCCGGCTAGCCTCTCAGCCAGGAAGCTTCCCTACAGAGCGGCCAGCA 240
DB 232 CTGCTCCAGCCGGCTAGCCTCTCAGCCAGGAAGCTTCCCTACAGAGCGGCCAGCA 291
QY 241 GGAAGCTATCTGAGGCGAGGCTGGGCCCTTATGCCAGGGGGCTGCCAGCCACATCTCC 300
DB 292 GGAAGCTATCTGAGGCGAGGCTGGGCCCTTATGCCAGGGGGCTGCCAGCCACATCTCC 351
QY 301 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGCAGAG 360
DB 352 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGCAGAG 411
QY 361 GACTGGTGCAGTGAACCGTACAGCTCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGT 420
DB 412 GACTGGTGCAGTGAACCGTACAGCTCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGT 471
QY 421 GTGGTGAAGCTGGCCCTACACGAAAGTGAACAGACAGACACTATGCAATGAAAGTCTTTC 480
DB 472 GTGGTGAAGCTGGCCCTACACGAAAGTGAACAGACAGACACTATGCAATGAAAGTCTTTC 531
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGTCCTCCAG 540
DB 532 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGTCCTCCAG 591
QY 541 GCTGCCAGGGAGCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACAGAGGATTT 600
DB 592 GCTGCCAGGGAGCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACAGAGGATTT 651
QY 601 GCCATCTGAAGAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660
DB 652 GCCATCTGNAGAGCTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 711
QY 661 CCAGCTGAGGACAACTCTATTG- 684
DB 712 CCGGCTGAGGACAACTCTATTG- 771
QY 685 ----- 684
DB 772 TCAACAATAATGCCAAGTCCCACTCCCTGCTTCCCTCTGAGCAGCAGAGAGTGGATCC 831
QY 685 -----GTGTTTGACCTCTCAGAAAGGGGCCGCTCATGGAGTCCCG 726
DB 832 ACGTGGGCTGGGCTCAGTGTGACCTCTGAGAAAGGGGCCGCTCATGGAGTCCCG 891
QY 727 TGTGACAAGCCCTTCTCGSAGGAGCAAGCTGCGCTCTACCTGGGGAGCTCATCTCTGGGC 786

DB 892 TCTGACAAGCCCTTCTCGAGGAGCAAGCTCGCTTACCTCGGGAGCTCATCTCTGGGC 951
QY 787 CTCGAGTACTTCACTTCCAGAGATCTCTCCAGAGGACATCAAGCCATCAACCTGCTC 846
DB 952 CTCGAGTACTTCACTTCCAGAGATCTCTCCAGAGGACATCAAGCCATCAACCTGCTC 1011
QY 847 CTGGGGGATGATGGGCACGTGAAGATCGCGACTTTTGGCGTCCAGCAACCAAGTTTGA 906
DB 1012 CTGGGGGATGATGGGCACGTGAAGATCGCGACTTTTGGCGTCCAGCAACCAAGTTTGA 1071
QY 907 AACAGCGCTCAGCTGTCCAGACGCGGGGAAACCCAGCAATTCATGSCCCCGAGGCAATT 966
DB 1072 AACAGCGCTCAGCTGTCCAGACGCGGGGAAACCCAGCAATTCATGSCCCCGAGGCAATT 1131
QY 967 TCTGATTCGGGCGCAGAGCTTCACTGGGAAGGCTTGGATGTATGGGCCACTGCGCTCACG 1026
DB 1132 TCTGATTCGGGCGCAGAGCTTCACTGGGAAGGCTTGGATGTATGGGCCACTGCGCTCACG 1191
QY 1027 TTGTACTGCTTTTCTATGGGAAGTSCCAATTCATCGACGATTTTCATCTGCGCCCTCCAC 1086
DB 1192 TTGTACTGCTTTTCTATGGGAAGTSCCAATTCATCGACGATTTTCATCTGCGCCCTCCAC 1251
QY 1087 AGAAGATCAAGAAATGAGCCCGTGGTGTCTCTGAGAGCCAGAAATCAGCAGGAGCTC 1146
DB 1252 AGAAGATCAAGAAATGAGCCCGTGGTGTCTCTGAGAGGCGCAGAAATCAGCAGGAGCTC 1311
QY 1147 AAGGACCTCATCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTCGGGTGCCAGAC 1206
DB 1312 AAGGACCTCATCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTCGGGTGCCAGAC 1371
QY 1207 ATCAAGTTGACCCCTTGGGTGACCAAGAAACGGGGAGAGCCCTTTCCTTCGAGGAGGAG 1266
DB 1372 ATCAAGTTGACCCCTTGGGTGACCAAGAAACGGGGAGAGCCCTTTCCTTCGAGGAGGAG 1431
QY 1267 CACTGAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAAGCTCATCCCC 1326
DB 1432 CACTGAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAAGCTCATCCCC 1491
QY 1327 AGCTGACACCGTGTATCTGGTGAAGTCCATGCTGAGGAAGGCTTCCTTTGGGAACCCG 1386
DB 1492 AGCTGACACCGTGTATCTGGTGAAGTCCATGCTGAGGAAGGCTTCCTTTGGGAACCCG 1551
QY 1387 TTTGAGCCCCCAGCACGGAGGGAAGCGCATCATCTGCTCCAGAAACCTTACTGGTG 1446
DB 1552 TTTGAGCCCCCAGCACGGAGGGAAGCGCATCATCTGCTCCAGAAACCTTACTGGTG 1611
QY 1447 AAGAGAGGTTGGTGAAGGGGCAAGCCAGAGCTCCCGGCGTCCAGGAGAGCGAG 1506
DB 1612 AAGAGAGGTTGGTGAAGGGGCAAGCCAGAGCTCCCGGCGTCCAGGAGAGCGAG 1671
QY 1507 GCTGCATCC 1515
DB 1672 GCTGCATCC 1680
RESULT 10
ADB53308
ID ADB53308 standard; DNA; 3411 BP.
XX
AC ADB53308;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3850.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN W02003065993-A2.

XX 14-AUG-2003.
XX
XX
XX 04-FEB-2003; 2003WO-US003482.
XX
XX 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3850; 874pp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 3411 BP; 823 A; 875 C; 968 G; 745 T; 0 U; 0 Other;
XX
XX Query Match 81.6%; Score 1236.6; DB 9; Length 3411;
XX Best Local Similarity 88.5%; Pred. No. 3.8e-298;
XX Matches 1341; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
XX
XX 1 ATGAGAGGGGTTCAGCTGTCTGTCAGAGATCTTCGGGAGAGCTGGTAGACGGGTG 60
XX 107 ATGAGAGGCGATCAGCGCTGTGTCAGACCTTCAGCGGAACTGGTAGCGGGTG 166
XX
XX 61 GCAGCCATCATGTGACTCTGAGGAGGACAGATGTGGCCAGAGCCCTACTAGAAC 120
XX 167 GCAGCCATCATGTGGCCACCTTAGAGAGCAGAGGGGCCAGAGCCCTGCCAGCAAT 226
XX
XX 121 GGTGTGAGACCCCAACACGGGCGAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
XX 227 GGTGTGAGACCCCAACCCCGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTGTTCAGA 286
XX
XX 181 CTGCTCCCAACCCGCTACCTCTGACCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 240
XX

Db 287 CCCACACAGTGGTCCCAGCCTCTCTGTAGAAAGTTCTCCTCGAGGAAAGACCAGCG 346
QY
241 GGAAGCTATCTGGAGCGCAGGCTTGGGCCCTTATGCCCAGGGGCCCTGCCACCATCTCTCC 300
Db
347 GGAAGCTGTCTAGAGGCTCAGGTTCGGCCCTTACTCTACAGGACCTGCCAGTACATGTCT 406
QY
301 CCCCGGGCTGGCGAGGCCACCAATCGAGTCCCAACAGCTGGCCCATCTCAGATGACAGAG 360
Db
407 CCTCGGGCTTGGCGGAGACCCACCATCGAGTCCCATGTGGCCATCTCAGACACAGAG 466
QY
361 GACTCGCTGACGTCAACCCAGTACAGCTGACAGATGAGATTGGCAAGGTCCTACCGT 420
Db
467 GACTGTGCAACTGAAACAGGTACAGCTGCAAGTGGAGTGGCAAGGTCCTATGAT 526
QY
421 GTGTGAGGCTGGCCTACAAAGAAAGTGAAGAAGACACACTATGCAATGAAAGTCTTTCC 480
Db
527 GTGTGAGGCTGGCCTCAATGAAAGGGAAGACACACACTATGCAATGAAAGTCTTTCC 586
QY
481 AAAAAGAAAGTACTGAAGCAGTATGGCTTCCAGCTGGCCCTCCCGAGAGGGTCCGAG 540
Db
587 AAAAAGAAAGTACTGAAGCAGTATGGCTTCCAGCTGGCCCTCCCGAGAGGGTCCCAA 646
QY
541 GCTGCCAGGAGGACACAGCAAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGAGATT 600
Db
647 GCTCTCAGGGAGGCCAGCCAAACAGCTGCTGCCCTCGAGCGGTGTATCCAGGAGATT 706
QY
601 GCCATCTGAAAGAGCTGGACCACTGTAATGTGGTCAAACCTGATCGAGTCTCTGGATGAC 660
Db
707 GCCATTTTAAAGAAACTGGACCACTGTAATGTAGTCAAGTTGATCGAGTCTCTGGATGAT 766
QY
661 CCAGCTGAGGACAACTCTATTGGTGTGTTGACCTCTCAGAAAGGGGCCCTCATGGA 720
Db
767 CTGCTGAAGCAATCTCTATTGGTGTGTTGACCTCTCAGAAAGGGACCACTCATGGA 826
QY
721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTCGGGAGCTCATC 780
Db
827 GTGCCCTCGACAAAGCCCTTCCAGAGGAGCAAGCTCGCCTCTACCTCGGGACATCATC 886
QY
781 CTGGCCCTCGAGTACTTGCATCTGCAGAGATGCTTCAAGGACATCAAGCCATCCAAAC 840
Db
887 CTGGCCCTCGAGTACTTGCATCTGCAGAGATGCTTCAAGGACATCAAGCCGTCCTCAAT 946
QY
841 CTGCTCTCTGGGGATGATGGCAGTGAAGTCCCGAGCTTGGGTCAGCAACACAGTTT 900
Db
947 CTGCTCTCTGGGAGATGGGCACGTGAAGTTCGCCACTTGGTGTGTCAGCAACACAGTTT 1006
QY
901 GAGGGAAACGACGCTCAGCTGTCCAGACGCGGGAAACCCAGCATTCATGGCCCCGAG 960
Db
1007 GAGGGAAATGATCTCAGCTGTCCAGTACGGCAGGACCCAGCATTCATGGCCCCGAG 1066
QY
961 GCCATTTCTGATTCGGCCAGAGCTTCAAGTGGGAGGCTTGGATGTATGGCCACATGGC 1020
Db
1067 GCCATCTCTGACACCGGCCAGAGCTTCAGTGGGAAAGGCTTGGATGTATGGCCACATGGG 1126
QY
1021 GTACAGTGTACTCTTGTCTATGGGAAGTGCCTTTCATCGACGATTTTCATCTCTGCC 1080
Db
1127 GTTACATGTATGTTTGTGCTACGGAAGTGCCTTTCATGTAGTATCATCTCTGCC 1186
QY
1081 CTCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAAATCAGCGAG 1140
Db
1187 CTGCACAGAAAGATCAAGAAATGAGCCCGTGGTGTTCCTCGAGGAGCCGAGGTTCAGCGAG 1246
QY
1141 GAGCTCAGGACCTGATCTTGAAGATGTTAGACAGAAATCCCGAGAGAGAAATTTGGGGTG 1200
Db
1247 GAATCAAGACCTGATCTCTGAAGATCTAGACAGAAATCCTTGAAGAAATTTGGGGTG 1306
QY
1201 CCAGACATCAAGTTCACCTTGGGTGACCAAGAAACGGGAGGAGCCCTTCTCTTCGGAG 1260
Db
1307 TCTGATATCAAGTTCACCTTGGGTGACCAAGCATTCGAGAGGAGCCCTCTCTTCAGAG 1366
QY
1261 GAGAGACATTCGACGCTGTGGAGGTGACAGAGGGGAGGTAAAGAACTCATGTCAGGCTC 1320
Db
1367 GAGAACACTGCAGTGTGGTAGAGGTGACTCGAGGAGGAGGTGAAGAACTCAGTCAAGCTC 1426

QY 1321 ATCCCAGCTGGACACCGGTGATCCTGGTGAAGTCCATGCTGAGGAAGGCTTCCTTTGGG 1380
Db |||||
QY 1427 ATCCCAGCTGGACACCGGTGATCCTGGTGAAGTCCATGCTGAGGAAGGCTTCCTTTGGG 1486
Db |||||
QY 1381 AACCCCTTTGAGCCCCAGGACCGAGGGAAGAGCGATCCATGCTGCTCCAGGAACCTTA 1440
Db |||||
QY 1487 AACCCATTTGAGCCCCAAGCACGAGGGAAGAAAGATCTATGCTGCGCCAGGAACTTA 1546
Db |||||
QY 1441 CTGGTGAAGAAAGGTTTGTGAAGGGGCAAGAGCCACAGACTCCCGCGTCCAGGAA 1500
Db |||||
QY 1547 CTGTTGAAGAAAGGATGTGGAAGAGGGGGCAAGAGCCACAGACTTCCCGAGTCCAGGAA 1606
Db |||||
QY 1501 GACGAGGCTGCATCC 1515
Db |||||
QY 1607 GATGAGGCTGCATCC 1621
Db |||||

RESULT 11

AAZ29223

ID AAZ29223 standard; cDNA; 2545 BP.

XX

AC AAZ29223;

XX

DT 28-FEB-2000 (first entry)

XX

XX Human cell signalling protein-2 encoding cDNA.

DE

XX Cell signalling protein-2; CSIGP-2; cell proliferation; arteriosclerosis;

KW

KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;

KW

KW Addison's disease; multiple sclerosis; ss.

XX

OS *Homo sapiens*.

XX

XX Key Location/Qualifiers

FH CDS

FT 241..1860

FT /*tag= a

FT /product= "Cell Signalling Protein-2"

XX

XX WO9558558-A2.

XX

XX 18-NOV-1999.

XX

XX 13-MAY-1999; 99WO-US010567.

XX

XX 13-MAY-1998; 98US-0085343P.

XX

XX 26-AUG-1998; 98US-0098010P.

XX

XX (INCY-) INCYTE PHARM INC.

XX

XX Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;

XX

XX Baughn MR, Yang J;

XX

XX WPI; 2000-086432/07.

XX

XX P-PSDB; AAY44239.

XX

XX Human cell signaling proteins useful for, e.g. diagnosing cell

XX

XX proliferative and inflammatory disorders.

XX

XX Claim 9; Page 80-81; 90pp; English.

XX

XX The present sequence is a cDNA obtained from Incyte clone 640521 of

CC

CC BRSTN03 library. It encodes cell signalling protein-2 (CSIGP-2). It is

CC

CC expressed in reproductive, nervous and developmental tissues. Fragments

CC

CC of CSIGP encoding nucleic acid can be used as hybridisation probe for

CC

CC detecting CSIGP related sequences or allelic variants. Recombinant CSIGP

CC can be produced in host cells by transforming them with genetically

CC engineered vectors. Agonists or antagonists can be used in the treatment

CC of cell proliferative and inflammatory disorders associated with

CC decreased or increased CSIGP expression. CSIGP is used in the diagnosis,

CC prevention and treatment of cell proliferative disorders like

CC arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory disorders

CC like AIDS, Addison's disease, multiple sclerosis, etc

XX

SQ Sequence 2545 BP; 542 A; 737 C; 760 G; 506 T; 0 U; 0 Other;

XX

XX Query Match 43.6%; Score 660.2; DB 3; Length 2545;

XX

XX Best Local Similarity 71.3%; Pred. No. 2.3e-154;

XX

XX Matches 885; Conservative 0; Mismatches 353; Indels 3; Gaps 1;

QY

267 GCCTTATGCCACGGGCGCTGCCAGCCACATCTCCCCCGGGCCTGGCGGAGGCCACCAT 326

Db

618 GCCCTACTCACCGTCAGCTCCCGCAGTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 677

QY

327 CAGTCCCAACAGTGGCCATCTCAGATGAGAGGAGTGGCTGACCTGAACAGGTACAA 386

Db

678 GGAGTCTCACCGCTCTCCATCAGGGGTATGCGAGGACTGTGTGACGTGAATCAGTATAC 737

QY

387 GCTGCAGAGTGGATTGGCAAGGTCCTACGCTGTGGTGGAGCTGGCTTACAAACGAAG 446

Db

738 CTTGAAGGATGAATTTGGAAGGGCTCTATGCTGCTCAAGTTGGCTTACATGAAA 797

QY

447 TGAAGACAGACACTATGCAATGAAAGTCTCTTTCCAAAAAGAGTTACTGAAGCAGTATGG 506

Db

798 TGACAATACCTACTATGCAATGAAGGTGCTGTCCAAAAAGAGTCTGATCCGGCAGCGCG 857

QY

507 CTTTCCAGTCCGCTCCCGGAGAGGTCCTCCAGGCTGCCAGGAGGAGGAGGAGGAGGAGG 566

Db

858 CTTTCCAGTCCGCTCCCGGAGGTCCTCCAGGCTGCCAGGAGGAGGAGGAGGAGGAGG 917

QY

567 GCTGCTGCCCTGGAGCGGTGTACCAGGAGATTGCCATCTCGAAGAGCTGACACACGT 626

Db

918 CAGGGGCCCATTTGACAGAGTGTACAGGAATTCATCTCCAGAGCTTGACCAACCCC 977

QY

627 GAATGTGTCAAACTGATCGAGTCTCTGGATGACCCAGCTGAGGAGCAACCTTATTGGT 686

Db

978 CAAATGTGTGAAGCTGCTGGAGTCTCGATGACCCCAATGAGGACCATCTGTACATGT 1037

QY

687 GTTTGACCTCTGAGAAAGGGGCCCTCATGGAAGTGCCTGTGACAGCCCTTTCGGA 746

Db

1038 GTTCGAACCTGGTCAACCAAGGGCCCGTGATGGAAGTSCCCACCTCAAAACCACTCTCTGA 1097

QY

747 GGAGCAAGCTCGCTCTACTCGGGAGCGTCATCTCTGGGCTCGAGTACTTTCACCTGCGCA 806

Db

1098 AGACCAAGGCCGCTTCTACTTCCAGGATCTGATCAAGGATCGAGTACTTACCTACCA 1157

QY

807 GAAGATCGTCCACAGGGAGCATCAAGCCATCCAACTGCTCTCTGGGGATGATGGCACGT 866

Db

1158 GAAGATCATCCACCGTGACATCAAACTTCCAACTCTCTGTCGGAGAAGATGGGCACAT 1217

QY

867 GAAGATCGCGGACTTTGGGCTCAGCAACAGTTTGAAGGGAACGACGCTCAGCTGTCCAG 926

Db

1218 CAAAGATCGCTGACTTTGGTGTAGCAATGAATTCAAAGGGCAGTGACGCGCTCTCTCCAA 1277

QY

927 CACGGGGGAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTCGGGCCAGAGCTT 986

Db

1278 CACCGTGGCAACGCCGCTTCTATGSCACCCGAGTCCGCTCTGAGACCCGCAAGATCTT 1337

QY

987 CAGTGGGAAGGCTTGGATGATATGGGCCACTGGCGTCACTGTGTACTGCTTTGTCTANGG 1046

Db

1338 CTCCTGGGAAGGCTTGGATGATGTTGGGCTATGGGCTGACACTATCTGCTTTGTCTTTGG 1397

QY

1047 GAAGTGGCCCATTCATGACGATTTTCATCTCTGGCCCTCCACAGGAAGATCAAGATGAGCC 1106

Db

1398 CCAAGTCCCATTCATGAGCAGCGGATCATGTGTTTACACAGTAAGATCAAGAGTCAGGC 1457

QY

1107 CGTGGTGTTCCTGAGGAGCGCAGAAATCAGCAGGAGCTCAAGGACCTGATCTCTGAAGAT 1166

Db

1458 CTTGGAATTTCCAGACCGCCGACATAGCTGAGGACTTGAAGSACCTGATACCCGTAT 1517

QY

1167 GTTAGCAAGAAATCCCGAGACGAAATTTGGGGTSCCAGACATCAAGTTGCACCCCTTGGGT 1226

Db

1518 GCTGGCAAGAACCCCGAGTCCGAGGATCGTGTGTGCGGAAATCAAGCTGCACCCCTGGGT 1577

QY

1227 GACCAAGACGGGAGGAGGCCCTTCTCTCGAGGAGGAGCAGCTGCAGCGCTGTGGAGGT 1286

XX

|||

|||

Db 1578 CACGAGGATGGGGCGGAGCGCTTGCGCTGGAGGATGAGAACTGCACGCTGGTCGAAGT 1637
QY 1287 GACAGAGGGGAGGTTAAGAACTCAGTCAGGCTATCCCGAGCTGAGCACGCGTGATCCT 1346
Db 1638 GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACACATTCGCCAGCTTGGCAACCGTGATCCT 1697
QY 1347 GGTGAAGTCCATGCTCAGGAAGGTTCTTTTGGAAACCGGTTTGAGCCCCAGCGCGGAG 1406
Db 1698 GGTGAAGACCATGATACGTAAAGCTCTTTTGGAAACCATTCGAGGGC---AGCGCG 1754
QY 1407 GGAAGAGCATCATGCTCGCTCCAGGAAACCTACTGTGAAAGAGGGTTTGGTGAAGG 1466
Db 1755 CGAGGAACGCTCACTGTACGCGCTGGAACCTTGCTCAAGAGCAGCGAGCAAGCA 1814
QY 1467 GGGCAGAGCCAGAGCTCCCGGGCTCCAGGAAGCAGG 1507
Db 1815 CCTCAGGGCACCGACCCCGCCCGGTGGGGAGGAGG 1855

RESULT 12
ADB75227
ID ADB75227 standard; cDNA; 4942 BP.
XX ADB75227;
AC
AT
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker cDNA.
XX
KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
XX
OS Homo sapiens.
XX
PN WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX
DR WPI; 2003-248033/24.
XX
PT New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
PS Claim 1; SEQ ID NO 51; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 4942 BP; 1169 A; 1294 C; 1356 G; 1122 T; 0 U; 1 Other;
Query Match 43.0%; Score 651.2; DB 9; Length 4942;
Best Local Similarity 71.3%; Pred. No. 5e-152;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;
QY 267 GCCTTATGCCACGGGGCCCTGCAGCCACATCTCCCCCGGGGCTGCGGAGGCCACCAT 326
Db 536 GCCCTACTACCCCGTCAGCTCCCGCAGTCCTCGCTCGGCTGCCCGGGCGGACAGT 595
QY 327 CGAGTCCCAACACGCTGGCCATCTCAGATGCAGAGGACTGCGTGACAGTCAACACAGTACAA 386
Db 596 GAGTCTCACCACGCTCCATCACGGTATGCAGACCTGTGTGACGTGAATCAGTATAC 655
QY 387 GCTCAGAGTGAATTTGGCAAGGTGCTTACCGTGTGTGAGGCTGGCTTACAAAGAAAG 446
Db 656 CCTGAAGGATGAAATTGGAAAGGGCTCTATGGTGTGCTCAAGTTGGCTTACAATGAAA 715
QY 447 TGAACAGACACACTATGCAATGAAAGTCTTTTCCAAAGAGAGTTACTGAACAGTATGG 506
Db 716 TGACAATACCTACTATGCAATGAGGTGCTGTCCAAAAGAGCTGATCCGGACGCGCG 775
QY 507 CTTTCCAGCTCGCCCTCCCGCAGAGGGTCCCGAGCTGCCAGGAGGACCAAGCA 566
Db 776 CTTTCCAGCTCGCCCTCCCGCAGGACCCCGCAGCTCCTGGAGGCTGCATCCAGCC 835
QY 567 GCTGCTGCCCTGGAGCGGGTGTACAGAGATTCGCATCTCTGAAGAGCTGGACACGT 626
Db 836 CAGGGCCCCCATTTGAGCAGGTGTACAGGAAATGCCATCTCTCAAGAAAGCTGGACACC 895
QY 627 GAATGTGTCAAACCTGATCGAGGCTCCTGGATGACCCAGCTGAGGACCACTTATTTGGT 686
Db 896 CAATGTGTGAAGCTGGTGGAGGCTCCTGGATGACCCCAATGAGGACCATCTGTACATGGT 955
QY 687 GTTTGACCTCTCAGAAAAGGGGCCGCTCATGGAAGTGCCCTGTGACAAAGCCCTTCTCGGA 746
Db 956 GTTCGAACCTGGTCAACCAAGGGCCGCTGATGGAAGTGCCCACTCAACCACTCTCTGA 1015
QY 747 GGAGCAAGCTCGCTCTACCTCGGGACGTCATCTCGGCTCGAGTCTTGCACTGGCA 806
Db 1016 AGACCAGGCCGCTTCTACTTCCAGGATCTGATCAAGGCGATCGAGTACTTACACTACCA 1075
QY 807 GAAGATCTGCACAGGGACATCAAGCCATCAACCTGCTCCTGGGGGATGATGGCAGCT 866
Db 1076 GAAGATCATCCCGTGACATCAAACTTCAACCTCCTGGTCGGAAGATGGGCACAT 1135
QY 867 GAAGATCGCCGACTTTGGCGTCAGCAACAGTTTGGGGGAAACGACGCTCAGCTCCAG 926
Db 1136 CAAGATCGTGCATTTGGTGTGAGCAATGAATTCAGGGGAGTGACGCGCTCTCTCCAA 1195
QY 927 CACGGGGAAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTCGGGCCAGAGTT 986
Db 1196 CACGCTGGGACGCGCGCGCTTCTATGGCACCGGAGTCGCTCTCTGAGAGCCCGAAGTCTT 1255
QY 987 CAGTGGGAAGGCTTTGGATGTATGGCCACTGGCGTCACTTGTACTCTTTGTCTATGG 1046
Db 1256 CTCTGGGAGGCTTTGATGTTTGGCCATGGGTGTGACATATACTGCTTTGCTTTGG 1315
QY 1047 GAATGCCCATTCATCGACGATTTTCACTCTGGCCCTCCACAGGAAGATCAAGAATGAGCC 1106
Db 1316 CCAATGGCCATTCATGACGAGCGGATCATGTGTTTACACAGTAAGATCAAGAGTCAGGC 1375
QY 1107 CGTGTGTTTCTCAGGAGCCAGAAATCAGCGAGGAGCTCAAGACCTGATCTCTGAAGAT 1166
Db 1376 CTTGGAATTTCCAGCACGACCCGACATAGCTGAGGACTTGAAGGACCTGATCAACCCGAT 1435
QY 1167 GTTAGCAAGAATCCCGAGAGAGAAATTTGGGGTCCACAGATCAAGTTGCACTTGGGT 1226
Db 1436 GCTGCAAGAAGACCCCGAGTCGAGATCGTGTGTCGGAAATCAAGCTGCAACCTTGGGT 1495
QY 1227 GACCAAGAACGGGAGGAGCCCTTCTTCCGAGAGGAGACCTGCGAGCGTGGTGGAGGT 1286
Db 1496 CACGAGGCATGGGGCGGAGCGCTTGGCGTCCGAGATGAGAACTGCACGCTGGTCCAGT 1555

QY 1167 GTTAGACAGATCCGACGAGATTTGGGTSCCAGACATCAAGTTCACCTTGGT 1226
Db 930 GTTGACAGAAACCCGAGTCGAGATCGTGTGCGGAAATCAAGCTGCACCCCTGGT 989
QY 1227 GACCAAGAACGGGAGGAGCCCTTCCCTTCGGAGGAGGAGCACTGCAGCGTGTGAGT 1286
Db 990 CACGAGGCATGGGGGAGCGCTTGGCGTCGAGGATGAGACTGCACGCTGGTGAAGT 1049
QY 1287 GACAGAGGGAGGTAGAACTCAGTCAGGTCAATCCAGCTGCACACCGTGAATCCT 1346
Db 1050 GACTGAAGAGGAGGTGAGAACTCAGTCAAAACACATTCAGCTGGCAACCGTGATCCT 1109
QY 1347 GGTGAAGTCCATGCTGAGGAGCGTTCCTTTGGGAACCGTTTGAGCCCGCAGGAG 1406
Db 1110 GGTGAAGACCATGATACGTAAACGCTCCTTTGGGAACCGTTTGAGGGG---AGCCGGC 1166
QY 1407 GGAAGAGCATCCATGCTGCTCCAGGAACCTACTGTGAAAGAGGTTTGGTGAAG 1466
Db 1167 GGAGAACGCTCACTGTCAGGCGCTGGAACTTGCTCACCAAAAACCAACGAGGAATG 1226
QY 1467 GGGCAAGAGCCAGAGCTC 1485
Db 1227 TGAGTCCCTGCTGAGCTC 1245

RESULT 14
ADB75229
ID ADB75229 standard; cDNA; 1804 BP.
AC ADB75229;
DT 04-DEC-2003 (first entry)
DE Prostate cancer marker cDNA.
XX Prostate; cancer; cytostatic; gene therapy; marker; ss.
KW Homo sapiens.
OS WO2003009814-A2.
PN 06-FEB-2003.
PD 25-JUL-2002; 2002WO-US023913.
PF 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362159P.
XX (MILL-) MILLENNIUM PHARM INC.
XX

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersht S, Kamatkar S, Wonsey AM, Glatk K, Zhao X, Anderson D;
XX WPI; 2003-248033/24.

XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.

XX Disclosure; SEQ ID NO 53; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1804 BP; 445 A; 519 C; 453 G; 387 T; 0 U; 0 Other;

QY Query Match 43.0%; Score 651; DB 9; Length 1804;
Best Local Similarity 71.5%; Pred. No. 3.9e-152;
Matches 871; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

QY 267 GCCTTATGCGCAGGGGCTGCAGCCACATCTCCCGCGGGCTGGGAGGCCACCAT 326

Db 30 GCGCTACTACCGCGTCACTCCCGCAGTCTCGCTCGCTGCGCTGCGCGCGCGACAGT 89

QY 327 CGAGTCCACACACGTCGCGCATCTCAGATGCAGAGGACTCGGTGACGCTGAACCACTACAA 386

Db 90 GGAGTCTCACACGCTCTCATCACGGGTATGCAGGACTGTGTGCAGCTGAATCAGTATAC 149

QY 387 GCTGAGAGTGAATTTGGCAAGGGTGCCTACGGTGTGTGAGGCTGGCTTACACCAAG 446

Db 150 CTTGAAGGATGAAATTTGAAAGGGCTCTATGGTGTCTCAAGTTGGCTTACATGAAA 209

QY 447 TGAAGACAGACATGCAATGAAAGTCTTTCCAAAAGAGTCTACTGAAGCAGTATGG 506

Db 210 TGCAATACCTACTATGCAATGAAAGTGTCTTCCAAAAGAGTGTATCCGGCAGCGCGG 269

QY 507 CTTTCCAGTCGCGCTCCCGGAGAGGCTCCAGGCTCCCGAGGAGGACCAAGCA 566

Db 270 CTTTCCAGTCGCGCTCCCGGAGGACCGCGCAGCTCTCGAGGCTGCATCCAGCC 329

QY 567 GCTGCTGCCCTGGAGCGGGTGTACAGGAGTTCCTCATCTCTGAAGAAGCTGGACACGT 626

Db 330 CAGGGCGCCCATTTGAGCAGGTGTACAGGAAATGCCATCTCTCAAGAAGCTGGACACC 389

QY 627 GAATGTGTCAAACTGATCGAGGCTCTCGATGAGCCAGCTGAGGACACCTCTATTGGT 686

Db 390 CAATGTGTGAAGCTGGTGGAGGCTCTGGATGAGCCCAATGAGGACCATCTGTACATGT 449

QY 687 GTTTGACCTCTGAGAAAGGGGCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGA 746

Db 450 GTTCGAATGCTCAACCAAGGCGCGTGTGAGTGCACCCCTCAACCACTCTCTGA 509

QY 747 GGAGCAAGCTCGCTCTACCTGCGGAGCTCATCTCGGCGCTCGAGTACTTGCATGCCA 806

Db 510 AGACGAGCGCGCTTCTACTTCCAGGATCTGATCAAGGCGCATCGAGTACTTACATCCA 569

QY 807 GAAGATGTCACAGGAGACATCAAGCCATCAACCTGCTCTCGGGGATGATGGCAGCT 866

Db 570 GAAGATCATCCCGTGACATCAAACTTCAACCTCTCTGCTCGAGGAAGATGGGACAT 629

QY 867 GAAGATCGCGCACTTTGGCGCTCAGCAACCACTTTTGGGGGAACGACGCTAGCTGTCCAG 926

Db 630 CAAGATCGCTGACTTTGGTGTGAGCAATGAATTCAGGCGCAGTACCGCGCTCTCTCCA 689

QY 927 CACGCGGAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTCGGGCCAGAGCTT 986

Db 690 CACCGTGGGACGCGCGCTTCTATGGCACCGGAGTGGCTCTCTGAGACCGCGCAAGATCTT 749

QY 987 CAGTGGGAAGGCGCTTGGATGTATGGGCACCTGGCGTCACTGTGTACTGCTTTGTCTATGG 1046

Db 750 CTCTGGGAAGGCGCTTGGATGTTGGGCCATGGGTGTGACACTATCTACTTGTCTTTGG 809

QY 1047 GAAGTGGCCATTCATCGACGATTTTCACTCTGGGCCCTCCACAGGAAGATCAAGAATAGCC 1106

Db 810 CCACTGGCCATTCATGACGAGCGGATCATGTGTTTACACAGTAAAGTCAAGAGTCAGGC 869

QY 1107 CGTGGTGTCTTCAGGAGGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCCTCAAGAT 1166

Db 870 CTTGAAATTTCCAGACCGCGCGGACATAGCTGAGGACTTGAAGGACCTGATCAACCGGTAT 929


```

Db 1226 CTCGGGAAGGCTTGGATGTTTGGCCATGGGTGACACTATACTGCTTGTCTTTGG 1285
QY 1047 GAAGTCCCAATTTCATCGACGATTTTCATCCTGGCCCTCCACAGGAAGATCAAGAATGAGCC 1106
Db 1286 CCAGTGGCCCAATTTCATGGACGAGCGGATCATGTGTTTACACAGTAAAGATCAAGAGTCAGGC 1345
QY 1107 CGTGGTGTCTCTGAGGAGCCAGAATCAGCCAGGAGCTCAAGGACCTGATCCTGAAGAT 1166
Db 1346 COTGGAATTTCCAGACCAGCCCGACATAGCTGAGGACTTGAAGGACCTGATCACCCGTAT 1405
QY 1167 GTTAGACAAGATCCCGAGACGAGAATTGGGGTGCCAGACATCAAGTTGCACCCCTTGGGT 1226
Db 1406 GCTGACACAGAAACCCCGAGTCAGGATCGTGTGCGGAATCAAGCTGCACCCCTGSGT 1465
QY 1227 GACCAAGAACGGGAGGAGCCCTTCCTTCGGAGGAGGAGCAGCTGACGCGTGGTGGAGGT 1286
Db 1466 CACGAGGCATGGGGCGGAGCCGTTGCCGTGCGGAGGATGAGAACTGCACGCTGTCGAAGT 1525
QY 1287 GACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCGAGCTGGACACCGTGATCCT 1346
Db 1526 GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACACATTCCCAGCTTGGCAACCGTGATCCT 1585
QY 1347 GGTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGGAACCCGTTTGAAGCCCAAGGACGGAG 1406
Db 1586 GGTGAGACCATGATACGTAAACGCTCCTTTGGGAACCCATTTCGAGGGC--AGCCGGCG 1642
QY 1407 GGAAGAGCGATCCATGCTCTGCTCCAGGAAACCTACTGTGTAAGAAGAGGTTTGGTGAAGG 1466
Db 1643 GGAGGAACGCTCACTGTCAGCGCTGGAACCTTGCTCACCAAAAAACCAACGAGGGAATG 1702
QY 1467 GGGCAAGAGCCCGAGGCTC 1485
Db 1703 TGAGTCCCTGTCTGAGCTC 1721

```

Search completed: July 25, 2004, 07:15:13
 Job time : 902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 22:25:04 ; Search time 9534 Seconds
(without alignments)
6887.424 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atgagggggggtccagctgt.....aggaagacgaggctgcattcc 1515

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1515	100.0	2190	6	AR221281	AR221281 Sequence
2	1515	100.0	2190	6	AX406674	AX406674 Sequence
3	1511.8	99.8	3529	9	AF425301	AF425301 Homo sapi
4	1511.8	99.8	3536	9	AF425232	AF425232 Homo sapi
5	1511.8	99.8	3575	9	BC043487	BC043487 Homo sapi
6	1511.8	99.8	3583	9	HSM801550	AL136576 Homo sapi
7	1510.2	99.7	1937	6	AX746188	AX746188 Sequence
8	1490.6	98.4	1683	6	AX455763	AX455763 Sequence
9	1490.6	98.4	2711	6	AX455761	AX455761 Sequence
10	1458.8	96.3	1542	6	AX166519	AX166519 Sequence
11	1457.2	96.2	3501	6	AX179641	AX179641 Sequence
12	1321.2	87.2	2469	9	BC031647	BC031647 Homo sapi
13	1241.4	81.9	3435	10	BC017529	BC017529 Mus muscu
14	1238.2	81.7	3429	10	S83194	S83194 Ca2+/calmod
15	1236.6	81.6	3411	10	AB023658	AB023658 Rattus no
16	1235	81.5	3406	10	RATGCDP	LA2810 Rattus norv
17	1224.6	80.8	3431	10	AF117384	AF117384 Mus muscu
18	685	45.2	3898	9	AK128601	AK128601 Homo sapi
19	661.8	43.7	1827	9	AB081336	AB081336 Homo sapi
20	661.8	43.7	2011	9	BC026060	BC026060 Homo sapi
21	661.8	43.7	5324	9	AF321387	AF321387 Homo sapi
22	660.2	43.6	2545	6	BD22105	BD22105 Cell sign
23	658.4	43.5	1572	5	AB083001	AB083001 Xenopus l
24	657	43.4	1720	9	AK024748	AK024748 Homo sapi
25	651.2	43.0	4450	9	AB018330	AB018330 Homo sapi
26	651	43.0	1804	9	AF101264	AF101264 Homo sapi
27	651	43.0	2975	9	AF287631	AF287631 Homo sapi
28	651	43.0	5611	9	AF287630	AF287630 Homo sapi
29	648.2	42.8	2661	10	BC023103	BC023103 Mus muscu
30	641.4	42.3	1767	6	AX056421	AX056421 Sequence
31	641.4	42.3	2772	9	AF140507	AF140507 Homo sapi
32	633.2	41.8	2132	10	AB018081	AB018081 Rattus no
33	628.4	41.5	1866	10	AF453383	AF453383 Mus muscu
34	526	34.7	1698	9	AB081337	AB081337 Homo sapi
35	526	34.7	2602	9	AF321385	AF321385 Homo sapi
36	526	34.7	5195	9	AF321388	AF321388 Homo sapi
37	526	34.7	5238	9	AF321386	AF321386 Homo sapi
38	515.6	34.0	4763	10	AK122370	AK122370 Mus muscu
39	455.6	30.1	3814	9	HSM805370	AL834322 Homo sapi
40	376.6	24.9	1080	9	BC000318	BC000318 Homo sapi
41	361	23.8	29629	6	AX406676	AX406676 Sequence
42	361	23.8	29629	6	AX406676	AX406676 Sequence
43	361	23.8	158414	9	AC005940	AC005940 Homo sapi
44	354.4	23.4	432	6	AX775531	AX775531 Sequence
45	305.6	20.2	1159	9	AF091074	AF091074 Homo sapi

ALIGNMENTS

RESULT 1
AR221281 LOCUS AR221281 Sequence 1 from patent US 6426206, 2190 bp DNA linear PAT 26-SEP-2002
DEFINITION AR221281
ACCESSION AR221281
VERSION AR221281.1 GI:23328248
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2190)
AUTHORS Wei, M.-H., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6426206-A 1 30-JUL-2002;

FEATURES	source	Location/Qualifiers
		1. .2190
		/organism="unknown"
		/mol_type="genomic DNA"
ORIGIN		
	Query Match	100.0%; Score 1515; DB 6; Length 2190;
	Best local Similarity	100.0%; Pred. No. 0;
	Matches 1515; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGAGGGGGTCCAGCTGCTCTGTCGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
Db	173	ATGAGGGGGGTCCAGCTGCTCTGTCGCAGATCCTCGGCAGAGCTGGTAGAACGGGTG 232
Qy	61	GCAGCCATCATGTGACTCTACTTTGAGGAGGCAGATGGTGGCCCGAGAGCCTTACTAGAAAC 120
Db	233	GCAGCCATCATGTGACTCTACTTTGGAGAGGCAGATGGTGGCCCGAGAGCCTTACTAGAAAC 292
Qy	121	GGTGTGACCCCCACCACACGGGGCAGAGCTGCCTCTGTGATCCCTGCGAGTACTTCAAGA 180
Db	293	GGTGTGACCCCCACCACGGGGCAGAGCTGCCTCTGTGATCCCTGCGAGTACTTCAAGA 352
Qy	181	CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db	353	CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 412
Qy	241	GGAAGTACTTGAGGGCGAGGCTGGGCCCTTATGCCACGGGGCTCCAGCCACATCTCC 300
Db	413	GGAAGTACTCTGAGGGCGAGGCTGGGCCCTTATGCCACGGGGCTCCAGCCACATCTCC 472
Qy	301	CCCCGGGCTTGGGGAGGCCACCATCGAGTCCACCAAGTGGGCATCTCAGATGCAGAG 360
Db	473	CCCCGGGCTTGGCGAGGCCACCATCGAGTCCACCAAGTGGGCATCTCAGATGCAGAG 532
Qy	361	GACTGCGTCAGCTGAACCAAGTCAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTAGCGT 420
Db	533	GACTGCGTCAGCTGAACCAAGTCAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTAGCGT 592
Qy	421	GTGCTGAGGCTGCCCTACAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTCC 480
Db	593	GTGCTGAGGCTGCCCTACAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTCC 652
Qy	481	AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCGCCCTCCCCGAGAGGGTCCAG 540
Db	653	AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCGCCCTCCCCGAGAGGGTCCAG 712
Qy	541	GCTGCCAGGAGGACAGCCAGCAGCTGCTGCCCTGGAGCGGGTGTCACAGGAGATT 600
Db	713	GCTGCCAGGAGGACAGCCAGCAGCTGCTGCCCTGGAGCGGGTGTCACAGGAGATT 772
Qy	601	GCCATCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660
Db	773	GCCATCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 832
Qy	661	CCAGCTGAGACAACTCTATTGTTGTTTGA CTTCTCTGAAGAAAGGGGCCCGCTCATGAA 720
Db	833	CCAGCTGAGACAACTCTATTGTTGTTTGA CTTCTCTGAAGAAAGGGGCCCGCTCATGAA 892
Qy	721	GTGCCCTGTACAAGCCCTTCTCGGAGGAGCAAGCTGCCCTCTACCTGCCGAGCTCATC 780
Db	893	GTGCCCTGTACAAGCCCTTCTCGGAGGAGCAAGCTGCCCTCTACCTGCCGAGCTCATC 952
Qy	781	CTGGGCTCGAGTACTTGTGACTGCGCAGAGATCGTCCACAGGGACATCAAGCCATCCAAC 840
Db	953	CTGGGCTCGAGTACTTGTGACTGCGCAGAGATCGTCCACAGGGACATCAAGCCATCCAAC 1012
Qy	841	CTGCTCTGGGGATGATGGGCAGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 900
Db	1013	CTGCTCTGGGGATGATGGGCAGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 1072
Qy	901	GAGGGGAACGACTCAGCTGTCTCAGCACGGCGGGAAACCCACAGATTTCATGGCCCCGAG 960
Db	1073	GAGGGGAACGACTCAGCTGTCTCAGCACGGCGGGAAACCCACAGATTTCATGGCCCCGAG 1132

Qy	961	GCATTTC	GATTCCGGCCAGAC	CTTCAGTGGGAAGGCCCTTGATGTATGGGCGA	CTGGC	1020	
Db	1133	GCATTTC	GATTCCGGCCAGAC	CTTCAGTGGGAAGGCCCTTGATGTATGGGCGA	CTGGC	1192	
Qy	1021	GTCA	CGTGTCTGCTATGGGAAGTCCCA	TTCCATCGACGATTTCA	CTCTGGCC	1080	
Db	1193	GTCA	CGTGTCTGCTATGGGAAGTCCCA	TTCCATCGACGATTTCA	CTCTGGCC	1252	
Qy	1081	CTCC	CAGGAAGATCAAGAA	TGAGCCCGTGGTGTTCCTGAGGAGCCAGAAAT	TCAGCGAG	1140	
Db	1253	CTCC	CAGGAAGATCAAGAA	TGAGCCCGTGGTGTTCCTGAGGAGCCAGAAAT	TCAGCGAG	1312	
Qy	1141	GAG	CTAAGACCTGATCTCTGA	GATGTTAGACAAGAA	TCCCGAGACGAAAT	TGGGTG	1200
Db	1313	GAG	CTAAGACCTGATCTCTGA	GATGTTAGACAAGAA	TCCCGAGACGAAAT	TGGGTG	1372
Qy	1201	CCAG	CATCAAGTTGCAC	CCCTTGGGTGACCAAGAA	CGGGGAGGAGCCCTT	CTCTGGAG	1260
Db	1373	CCAG	CATCAAGTTGCAC	CCCTTGGGTGACCAAGAA	CGGGGAGGAGCCCTT	CTCTGGAG	1432
Qy	1261	GAG	GACATG	CAGCGTGTGAGGTGACAGAGGGGAGGTTAAGAA	CTCAGTCA	GCGCTC	1320
Db	1433	GAG	GACATG	CAGCGTGTGAGGTGACAGAGGGGAGGTTAAGAA	CTCAGTCA	GCGCTC	1492
Qy	1321	ATCC	CAGCTGGACAC	CGGTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTC	CTTTGGG	1380	
Db	1493	ATCC	CAGCTGGACAC	CGGTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTC	CTTTGGG	1552	
Qy	1381	AA	CCCGTTTGAG	CCCCCAGGCAGGAGGGAAGCGATCCATGCTCTCAGGAA	ACCTA	1440	
Db	1553	AA	CCCGTTTGAG	CCCCCAGGCAGGAGGGAAGCGATCCATGCTCTCAGGAA	ACCTA	1612	
Qy	1441	CTG	GTAAAGAGG	GTGTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGA		1500	
Db	1613	CTG	GTAAAGAGG	GTGTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGA		1672	
Qy	1501	GAC	GAGCTGCAT	CC	1515		
Db	1673	GAC	GAGCTGCAT	CC	1687		
RESULT 2							
LOCUS	AX406674						
DEFINITION	Sequence 1 from Patent WO224920.						
ACCESSION	AX406674						
VERSION	AX406674.1 GI:21439624						
KEYWORDS	Homo sapiens (human)						
SOURCE	Homo sapiens						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1						
AUTHORS	Beasley,E.M., Wei,M.H., Bonazzi,V.R., Sanders,R.C. and di Francesco,V.C.						
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof						
JOURNAL	Patent: WO 0224920-A 1 28-MAR-2002;						
FEATURES	PE Corporation (NY) (US)						
	Location/Qualifiers						
	1..2190						
source	/organism="Homo sapiens"						
ORIGIN	/mol_type="unassigned DNA"						
	/db_xref="taxon:9606"						
Query Match	100.0%; Score 1515; DB 6; Length 2190;						
Best Local Similarity	100.0%; Pred. No. 0;						
Matches 1515; Conservative	0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	ATG	GAGGGGGG	GTCTGCTCCAGATCTTCGGGCAGAGCTGGTAGAACGGTG	60		

Db 173 ATGAGGGGGGTCAGCTGTCTGTGCTGCCAGGATCTCTGGGCGAGAGCTGGTAGAACGGGTG 232
QY 61 GAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCGAGAGCTTACTAGAAAC 120
Db 233 GAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCGAGAGCTTACTAGAAAC 292
QY 121 GGTGTGGACCCGCCACACGAGGCGCAGAGCTGCTCTGTGATCCCTGGCGAGTACTTCAAGA 180
Db 293 GGTGTGGACCCGCCACACGAGGCGCAGAGCTGCTCTGTGATCCCTGGCGAGTACTTCAAGA 352
QY 181 CTGCTCCACAGCGCGCTAGCTCTCAGCCAGGAGAGCTTTCCTACAGAGGCGCCAGCA 240
Db 353 CTGCTCCACAGCGCGCTAGCTCTCAGCCAGGAGAGCTTTCCTACAGAGGCGCCAGCA 412
QY 241 GGAAGCTATCTGAGGCGCAGAGCTGGGCCCTTATGCCACGGGGCCCTGCCAGCCACATCTCC 300
Db 413 GGAAGCTATCTGAGGCGCAGAGCTGGGCCCTTATGCCACGGGGCCCTGCCAGCCACATCTCC 472
QY 301 CCCGGGCTTGGCGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGAGAG 360
Db 473 CCCGGGCTTGGCGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGAGAG 532
QY 361 GACTGCGTGCAGCTGAACAGTACAGCTGAGAGTGGCAAGGCTGGCAAGGCTGCCATACGGT 420
Db 533 GACTGCGTGCAGCTGAACAGTACAGCTGAGAGTGGCAAGGCTGGCAAGGCTGCCATACGGT 592
QY 421 GTGTGAGGCTGGCCCTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 480
Db 593 GTGTGAGGCTGGCCCTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 652
QY 481 AAAAAGAGTTACTAAGCAGTATGGCTTCCAGTGGCCCTCCCGGAGAGGTCAG 540
Db 653 AAAAAGAGTTACTAAGCAGTATGGCTTCCAGTGGCCCTCCCGGAGAGGTCAG 712
QY 541 GCTGCCAGGAGGAGCAGCAGCTGCTGCCCTGGAGCGGCTGTACAGGAGATT 600
Db 713 GCTGCCAGGAGGAGCAGCAGCTGCTGCCCTGGAGCGGCTGTACAGGAGATT 772
QY 601 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACCTGATCGAGTCTCTGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACCTGATCGAGTCTCTGATGAC 832
QY 661 CCAGTGGAGCAACCTTATTTGTTGTTGATCTTCCCTGAGAAAGGCGCGTCAAGAA 720
Db 833 CCAGTGGAGCAACCTTATTTGTTGTTGATCTTCCCTGAGAAAGGCGCGTCAAGAA 892
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGAGCGTCAATC 780
Db 893 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGAGCGTCAATC 952
QY 781 CTGGGCTTCGAGTACTTGCATGCGAGAGATGCTCCACAGGACATCAAGCCATCCAAAC 840
Db 953 CTGGGCTTCGAGTACTTGCATGCGAGAGATGCTCCACAGGACATCAAGCCATCCAAAC 1012
QY 841 CTGCTCTCGGGGATGATGGGACGTGAGATCGCGACTTTGGCGTCAGCAACAGTTT 900
Db 1013 CTGCTCTCGGGGATGATGGGACGTGAGATCGCGACTTTGGCGTCAGCAACAGTTT 1072
QY 901 GAGGGAAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCGAG 960
Db 1073 GAGGGAAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCGAG 1132
QY 961 GCCATTTCTGATTCGGGCGCAGAGCTTCAAGTGGGAAGGCGCTTGGATGATGGGCCACTGGC 1020
Db 1133 GCCATTTCTGATTCGGGCGCAGAGCTTCAAGTGGGAAGGCGCTTGGATGATGGGCCACTGGC 1192
QY 1021 GTCAGTTGTACTGCTTTGTCTATGGGAAGTGGCCCATTCATCGAGATTCTATCTGGCC 1080
Db 1193 GTCAGTTGTACTGCTTTGTCTATGGGAAGTGGCCCATTCATCGAGATTCTATCTGGCC 1252
QY 1081 CTCACAGGAAGATCAAGAAATGAGCCCTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAAGATCAAGAAATGAGCCCTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1312

QY 1141 GAGCTCAAGGAGCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAATTTGGGTG 1200
Db 1313 GAGCTCAAGGAGCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAATTTGGGTG 1372
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTGGAG 1260
Db 1373 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTGGAG 1432
QY 1261 GAGGAGCACTGACAGCGTGGTGGAGTGCAGAGAGGGGAGGTTAAAGATCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTGACAGCGTGGTGGAGTGCAGAGAGGGGAGGTTAAAGATCAGTCAGGCTC 1492
QY 1321 ATCCCCAGCTGACACACGCTGATCTGTGAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1380
Db 1493 ATCCCCAGCTGACACACGCTGATCTGTGAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1552
QY 1381 AACCCGTTTGGAGCCCGAGGCGAGGAGGAAAGCGATCAATGCTGCTCCAGGAACCTTA 1440
Db 1553 AACCCGTTTGGAGCCCGAGGCGAGGAGGAAAGCGATCAATGCTGCTCCAGGAACCTTA 1612
QY 1441 CTGTGGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1500
Db 1613 CTGTGGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687

RESULT 3

AF425301
LOCUS Homo sapiens CAMKK alpha protein (CAMKK1) mRNA linear PRI 04-OCT-2002
DEFINITION
ACCESSION AF425301
VERSION AF425301.1 GI:23499315
KEYWORDS

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3529)
Harvey, M., Carra, S., Tascadda, F. and Barden, N.
Characterization of human CamKK alpha gene structure
Unpublished
2 (bases 1 to 3529)
Harvey, M., Carra, S., Tascadda, F. and Barden, N.
Direct Submission
Submitted (27-SEP-2001) Neuroscience, CHUL, 2705 Blvd. Laurier,
Ste-Foy, Quebec G1V4G2, Canada
Location/Qualifiers

REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1. 3529
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue_type="brain"
1. 3529
/gene="CAMKK1"
95.1612
/note="protein kinase"
/codon_start=1
/product="CAMKK alpha protein"
/protein_id="AAN37387.1"
/db_xref="GI:23499315"
/translation="MEGPAVCCQDPRAELVERVAAIDVTHLEADGGPEPTRNGVDP
PPRAAASVLPSTSLRLLPARPSLSARKLQBERPAGSYLEAOAGPYATGPASHISPR
AWRPPTTESHHVAISDAEDCVQNOYKLOSEIGKAYGVYRLVLA NESEDRIYAMKVL
KKLLKQYGFRRPPRPGSAQKQLPLPLERVYQETAILKKLDHVNKKLIEVL
DDPAEDNLXILVFDLLRGPVMEPCDKPFSEBQARLYLRDVLGLFYLHCQKIVHRDI
KPSNLLIGDDGHVKIADFGVSNFEGNDQLSSSTAGTAPFAPMAPEAISDSQGSFSGKAL

gene

CDS

DVWATGVTLYCFVYKCPFIIDFLLALHRIKNEPVVFPPEPISBELKDLILKMLDK
NPTRIGVDPIDKLPWTKNGEPLPSEEEHCSVTEVEEVKNSVRLI PSMWTVILV
KSMRLKRSFGNPPFPQARRERSMSAPGNLLVKEFGEGGKSPELPGVQDEAAS"

ORIGIN

Query Match 99.8%; Score 1511.8; DB 9; Length 3529;
Best local similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	ATGAGAGGGGGTCCAGCTGTCTGTCGCCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	60
Db	95	ATGAGAGGGGGTCCAGCTGTCTGTCGCCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	154
Qy	61	GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC	120
Db	155	GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC	214
Qy	121	GGTGTGACCCCCACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	180
Db	215	GGTGTGACCCCCACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	274
Qy	181	CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	240
Db	275	CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	334
Qy	241	GGAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGGGGCTGCCAGCACATCTCC	300
Db	335	GGAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGGGGCTGCCAGCACATCTCC	394
Qy	301	CCCCGGGCTGGGGAGGCCACCATGAGTCCACCAAGCTGGGCTGAGTGCAGTGCAGAG	360
Db	395	CCCCGGGCTGGGGAGGCCACCATGAGTCCACCAAGCTGGGCTGAGTGCAGTGCAGAG	454
Qy	361	GACTGCGTGCAGCTGAACAGCTACAAGCTGAGAGTGGCAAGGGTGGCTACGGT	420
Db	455	GACTGCGTGCAGCTGAACAGCTACAAGCTGAGAGTGGCAAGGGTGGCTACGGT	514
Qy	421	GTGTGAGGCTGGGCTTACAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC	480
Db	515	GTGTGAGGCTGGGCTTACAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC	574
Qy	481	AAAGAGAGTACTCAAGCAGTATGGCTTCCAGCTGCGCTCCCGGAGAGGTCCTCAG	540
Db	575	AAAGAGAGTACTCAAGCAGTATGGCTTCCAGCTGCGCTCCCGGAGAGGTCCTCAG	634
Qy	541	GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT	600
Db	635	GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT	694
Qy	601	GCCATCTGAAGAAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGTCTCGATGAC	660
Db	695	GCCATCTGAAGAAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGTCTCGATGAC	754
Qy	661	CCAGCTGAGGACCACTCTATTGGTGTTCACCTCTGAGAAAGGGCCCGTTCATGAA	720
Db	755	CCAGCTGAGGACCACTCTATTGGTGTTCACCTCTGAGAAAGGGCCCGTTCATGAA	814
Qy	721	GTGCCCTGTGACAAAGCCCTCTCGGAGGAGCAAGCTCGCCCTTACCTCGGGAGCGTCATC	780
Db	815	GTGCCCTGTGACAAAGCCCTCTCGGAGGAGCAAGCTCGCCCTTACCTCGGGAGCGTCATC	874
Qy	781	GTGGGCTCGAGTACTTGCATGTCAGNAGATGCTCCACAGGAGCATCAAGCCATCCAC	840
Db	875	GTGGGCTCGAGTACTTGCATGTCAGNAGATGCTCCACAGGAGCATCAAGCCATCCAC	934
Qy	841	CTGCTCTGGGGGATGATGGCAGCTGAAGATCGCCAGCTTTGGCGTCAGCAACAGATT	900
Db	935	CTGCTCTGGGGGATGATGGCAGCTGAAGATCGCCAGCTTTGGCGTCAGCAACAGATT	994
Qy	901	GAGGGGAACGAGCTGTGAGTGTCCAGACGCGCGGGAACCCGAGCATTCATGGCCCCGAG	960
Db	995	GAGGGGAACGAGCTGTGAGTGTCCAGACGCGCGGGAACCCGAGCATTCATGGCCCCGAG	1054

Qy	961	GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGSCCTTGGATGTATGGGCCACTGGC	1020
Db	1055	GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGSCCTTGGATGTATGGGCCACTGGC	1114
Qy	1021	GTACAGTGTACTGCTTTGTCTATGGGAGTGCCTTATTCATCGAGATTTTCATCTCGCC	1080
Db	1115	GTACAGTGTACTGCTTTGTCTATGGGAGTGCCTTATTCATCGAGATTTTCATCTCGCC	1174
Qy	1081	CTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTTCTTCCAGGAGCCAGAAATCAGCGAG	1140
Db	1175	CTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTTCTTCCAGGAGCCAGAAATCAGCGAG	1234
Qy	1141	GAGCTCAAGAACCTGATCTTGAAGATGTTAGACAAGAAATCCCAGAGAGGATTTGGGTG	1200
Db	1235	GAGCTCAAGAACCTGATCTTGAAGATGTTAGACAAGAAATCCCAGAGAGGATTTGGGTG	1294
Qy	1201	CCAGACATCAAGTTGCAACCTTGGGTGACCAAGAACCGGGAGGAGCCCTTCCCTTCGGAG	1260
Db	1295	CCAGACATCAAGTTGCAACCTTGGGTGACCAAGAACCGGGAGGAGCCCTTCCCTTCGGAG	1354
Qy	1261	GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC	1320
Db	1355	GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC	1414
Qy	1321	ATCCCGAGCTGGACCAAGCTGATCCTGTTGAAGTCCATGCTGAGGAGGCTTCTTTGGG	1380
Db	1415	ATCCCGAGCTGGACCAAGCTGATCCTGTTGAAGTCCATGCTGAGGAGGCTTCTTTGGG	1474
Qy	1381	AACCCGTTTGAAGCCCAAGGACCGAGGAAAGAGCGATCCATGCTGCTCCAGGAAACCTA	1440
Db	1475	AACCCGTTTGAAGCCCAAGGACCGAGGAAAGAGCGATCCATGCTGCTCCAGGAAACCTA	1534
Qy	1441	CTGTGAAAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGGAA	1500
Db	1535	CTGTGAAAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGGAA	1594
Qy	1501	GACGAGGCTGCATCC	1515
Db	1595	GACGAGGCTGCATCC	1609

RESULT 4

AF425232
LOCUS Homo sapiens CaMKK alpha protein mRNA, complete cds.
DEFINITION Homo sapiens CaMKK alpha protein mRNA, complete cds.
ACCESSION AF425232
VERSION AF425232.1 GI:23499313
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3536)
AUTHORS Tascenda, F., Carra, S., Harvey, M. and Barden, N.
TITLE Characterization of human CaMKK alpha gene structure
JOURNAL Unpublished

REFERENCE

2 (bases 1 to 3536)
AUTHORS Tascenda, F., Carra, S., Harvey, M. and Barden, N.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2001) Pharmacology, University of Modena and Reggio Emilia, Campi 183, Modena, MO 41100, Italy

FEATURES

Location/Qualifiers
1..3536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue_type="brain"

source

102..1619
/note="protein kinase kinase; CaMKK 1"
/codon_start=1
/product="CaMKK alpha protein"

CDS

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

JOURNAL
MEDLINE
PUBMED
22388257
12477932
2 (bases 1 to 3575)
Strausberg, R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb-x@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 89 Row: h Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27437009.
Location/Qualifiers
1..3575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:49844 IMAGE:5751570"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..3575
/gene="CAMKK1"
/note="synonyms: DKFZp761M0423, CAMKKA, MGC34095"
/db_xref="locusID:84254"
125..1642
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase 1
alpha, isoform a"
/protein_id="AAH43487.1"
/db_xref="GI:27694084"
/db_xref="locusID:84254"
translation="MEGGPAVCCQDPRAELVERVAALDVTHLEADGGPPTNRGVDP
PPRAASVTPGSTSRLLPAPRSLARKLSQERPAQSLYLAQAGPVPATPASHISPR
AWRRPTTESHHVALSDAEDCVQLNQKSEIGKAYGVVRLAYNESDRHYAMKVL
KKLLQKYGFPFRPPRSGSAQGGPAKQLPLERVYQETAILKLKLDHVVVKKLIEVL
DPAEDNLVLDLRKGPVMEVPCDFSEQARLYLRDVLGLVLEHCKI VHRDI
KPSNLILGDGHVADIGVSNQFEGNDAQLSSFAGTFAFWAPEAISDSGSGSKAL
DVWATGVLYCFYVGKCFDIDFILAHLRKLKNEPVPFPEPEISEKLLILKMLDK
NPETRIGVFDIPLKHPWTKNGEEBPLPSEEEHCSEVTEBEVKNSRLIIPSWTIVILV
KSLMRKRSFGNPFQFARREERSNAPGNLLVKEGFGEGGKSPGLPGVDEAAS"

FEATURES
source
misc_feature 506..1351
/note="S-TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"
ORIGIN
Query Match 99.8%; Score 1511.8; DB 9; Length 3575;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAGGGGGTCCAGCTGTCTGCTCCAGAGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 60
DB 125 ATGAGAGGGGGTCCAGCTGTCTGCTCCAGAGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 184
QY 61 GCAGCCATCATGTGACTACTTGGAGAGGAGATGTGTGCGCCAGAGCCCTACTAGAAC 120
DB 185 GCAGCCATCATGTGACTACTTGGAGAGGAGATGTGTGCGCCAGAGCCCTACTAGAAC 244
QY 121 GGTGTGACCCGCCACCACCGGGCCAGAGCTCTCTGTATCCCTGGCAGTACTTCAAGA 180
DB 245 GGTGTGACCCGCCACCACCGGGCCAGAGCTCTCTGTATCCCTGGCAGTACTTCAAGA 304
QY 181 CTGCTCCAGCCCGGCCCTAGCTCTCTAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
DB 305 CTGCTCCAGCCCGGCCCTAGCTCTCTAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 364
QY 241 GGAAGCTATCTGAGGCGGAGCTGGSCCTTATGCCAGGGGCGCTGCCAGCCATCTCC 300
DB 365 GGAAGCTATCTGAGGCGGAGCTGGSCCTTATGCCAGGGGCGCTGCCAGCCATCTCC 424
QY 301 CCCGGGCGTGGCGGAGGCCACCATCGAGTCCACCACCGTGGCCATCTCAGATGCGAGAG 360
DB 425 CCCGGGCGTGGCGGAGGCCACCATCGAGTCCACCACCGTGGCCATCTCAGATGCGAGAG 484
QY 361 GACTCGCTGAGCTGAACACAGTACAAAGCTGAGAGTGAAGTGGCAAGGGTGCCTACGGT 420
DB 485 GACTCGCTGAGCTGAACACAGTACAAAGCTGAGAGTGAAGTGGCAAGGGTGCCTACGGT 544
QY 421 GTGGTGGAGCTGGCTACAGAGAGTGAAGAGACAGACACTATGCAATGAAGTCTTTCC 480
DB 545 GTGGTGGAGCTGGCTACAGAGAGTGAAGAGACAGACACTATGCAATGAAGTCTTTCC 604
QY 481 AAAAAGAAAGTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGCGAGAGGTCCTCCAG 540
DB 605 AAAAAGAAAGTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGCGAGAGGTCCTCCAG 664
QY 541 GCTGCCCGAGGAGGACAGCCAGCAGCTGCTGCTGCCCTGGAGCGGGGTACAGAGAGATT 600
DB 665 GCTGCCCGAGGAGGACAGCCAGCAGCTGCTGCTGCCCTGGAGCGGGGTACAGAGAGATT 724
QY 601 GCCATCTGAGAGAGCTGGACCACTGTAATGTGTCAAACTGATCGAGGTCCTGGATGAC 660
DB 725 GCCATCTGAGAGAGCTGGACCACTGTAATGTGTCAAACTGATCGAGGTCCTGGATGAC 784
QY 661 CCAGCTGAGGACCACTCTATTTGGTGTGTTGACCTCTGAGAAAGGGGCCCGTCATGAA 720
DB 785 CCAGCTGAGGACCACTCTATTTGGTGTGTTGACCTCTGAGAAAGGGGCCCGTCATGAA 844
QY 721 GTGCCCTGTGAAGCCCTTCTGGAGGAGCAAGCTCGCTCTACCTCGGGAGCGTCATC 780
DB 845 GTGCCCTGTGAAGCCCTTCTGGAGGAGCAAGCTCGCTCTACCTCGGGAGCGTCATC 904
QY 781 GTGCCCTGTGAAGCCCTTCTGGAGGAGCAAGCTCGCTCTACCTCGGGAGCGTCATC 840
DB 905 GTGCCCTGTGAAGCCCTTCTGGAGGAGCAAGCTCGCTCTACCTCGGGAGCGTCATC 964
QY 841 CTGCTCTCGGGGAGTATGGGCAAGTGAAGATGCCGACTTTTGGCGTCAGCAACAGTTT 900
DB 965 CTGCTCTCGGGGAGTATGGGCAAGTGAAGATGCCGACTTTTGGCGTCAGCAACAGTTT 1024
QY 901 GAGGGGAACCAAGCTGAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 960
DB 1025 GAGGGGAACCAAGCTGAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 1084


```
QY 961 GCATTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATATGGGCACCTGGC 1020
|||
Db 1085 GCCATTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATATGGGCACCTGGC 1144
QY 1021 GTACGTTGTACTGCTTTGTCTATGGAAGTGCCCATTCATCAGCAATTCATCCTGGCC 1080
|||
Db 1145 GTCACGTTGTACTGCTTTGTCTATGGAAGTGCCCATTCATCAGCAATTCATCCTGGCC 1204
QY 1081 CTCACAGGAAGATCAAGATGATGCGCGTGTCTTCTGAGGAGCAGAAATCAGCGAG 1140
|||
Db 1205 CTCACAGGAAGATCAAGATGATGCGCGTGTCTTCTGAGGAGCAGAAATCAGCGAG 1264
QY 1141 GAGCTCAAGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1200
|||
Db 1265 GAGCTCAAGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1324
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAAACGGGAGAGGCCCTTCCTCGGAG 1260
|||
Db 1325 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAAACGGGAGAGGCCCTTCCTCGGAG 1384
QY 1261 GAGGACATCTGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
|||
Db 1385 GAGGACATCTGACGCTGTGGAGGTGACAGAGGAGGTTAAGAACTCAGTCAGGCTC 1444
QY 1321 ATCCCAGCTGACACCGTGTCTGTTGAAGTCCATGCTGAGGAAGCTTCCTTTGGG 1380
|||
Db 1445 ATCCCAGCTGACACCGTGTCTGTTGAAGTCCATGCTGAGGAAGCTTCCTTTGGG 1504
QY 1381 AACCGTTTGGCCCCAGACGAGGAGGAGGAGCGATCCATGCTGCTCAGGAAACCTTA 1440
|||
Db 1505 AACCGTTTGGCCCCAGACGAGGAGGAGGAGCGATCCATGCTGCTCAGGAAACCTTA 1564
QY 1441 CTGCTGAAGAAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGAA 1500
|||
Db 1565 CTGCTGAAGAAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGAA 1624
QY 1501 GACGAGGCTGCATCC 1515
|||
Db 1625 GACGAGGCTGCATCC 1639

RESULT 6
HSM801550 3583 bp mRNA linear PRI 10-MAR-2001
LOCUS Homo sapiens mRNA; cDNA DKFp761M0423 (from clone DKFp761M0423);
DEFINITION complete cds.
ACCESSION AL136576
VERSION AL136576.1 GI:13276654
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wiemann,S., Weil,B., Wellenreuther,R., Gassenhuber,J., Glassl,S.,
Ansorge,W., Boecker,H., Bloeker,H., Bauersachs,S., Blum,H.,
Lauber,J., Duesterhoeft,A., Beyer,A., Koehler,K., Strack,N.,
Mewes,H.W., Oostenwelder,B., Obermaier,B., Tampe,J., Heubner,D.,
Wambutt,R., Korn,B., Klein,M. and Poustka,A.
Toward a catalog of human genes and proteins: sequencing and
analysis of 500 novel complete protein coding human cDNAs
Genome Res. 11 (3), 422-435 (2001)
21154917
PUBMED 11230166
REFERENCE
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B. and
Wiemann,S.
Direct Submission
Submitted (18-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
JOURNAL
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
```

sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source
1..3583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFp761M0423"
/db_xref="taxon:9606"
/map="26.0 cR from top of Chr17 linkage group"
/clone="DKFp761M0423"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
1..3583
/gene="DKFp761M0423"
149..1666
/gene="DKFp761M0423"
/note="similarity to Cam-kinase alpha (Rattus norvegicus)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66511.1"
/db_xref="GI:13276655"
/db_xref="GOA:Q9BQH3"
/db_xref="SPTREMBL:Q9BQH3"
/translation="MEGPAVCCQPRAEILVERVAALDVTHLEADGGPEPTRNGVDP
PPRAASVTPGSTRLILPAPSLQSLARKLSQERFAGSYLAQAGPYATGPASHISPR
AWRRPTIESHHVAISDAECVQLNQKQSEIGKAGYGVRLAYNESEDRHYAMKVL
KKKLLKQVFPFRPPRPSQAAGGPAKOLLPLRYQETAILKLDHVNVLIVL
DDPAEDNLYLVFDLLRKGVMEVPCDKPSEEQARLYLRDVLGLELHCOKIVHRDI
KPSNLLIGDDGHVKIADFGVSNQFSGNQAOLSTAGTAPFAPEAISDSQSPSGKAL
DYVATGVTLCEVYKCKPIDDIFILALHRKIKNEPVVFEEPEISEELKDLILKMLDK
NPETRIGVPDIIKLFHWYTKNGBEPLPFSBEHCVSVEVTEEVKNSVRLILPISWTTVLV
KSMRLRRSPGNDLFEFPAARREERSMAGPNLLVKEGFGGKSPFLPGVQEDFAAS"
3525..3530
/gene="DKFp761M0423"
3546
/gene="DKFp761M0423"

polya_signal

polya_site

ORIGIN

Query Match 99.8%; Score 1511.8; DB 9; Length 3583;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTAGAACGGGTG 60
Db |||
149 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTAGAACGGGTG 208
QY 61 GCAGCCATCGATGTCACTCACTTGAGGAGGAGGAGAGCTGTGGCCAGAGCTACTAGAAC 120
Db 209 GCAGCCATCGATGTCACTCACTTGAGGAGGAGGAGAGCTGTGGCCAGAGCTACTAGAAC 268
QY 121 GGTGTGGACCCCGCCACCGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 269 GGTGTGGACCCCGCCACCGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 328
QY 181 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCTCAGAGAGGGCGCAGCA 240
Db 329 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTTTCCTCAGAGAGGGCGCAGCA 388
QY 241 GGAAGCTTCTGGAGGCGCAGCTGGGCGCTTATGCCAGGGGCGCTGCCAGCCACATCTCC 300
Db 389 GGAAGCTTCTGGAGGCGCAGCTGGGCGCTTATGCCAGGGGCGCTGCCAGCCACATCTCC 448
QY 301 CCCCGGCGCTGGCGGAGGCCACCATCGAGTCCCAACACCGTGGCCATCTCAGATCAGAG 360
|||

Db 449 CCCCCGGCTGGCGAGGCCACCATCGAGTCCACACAGCTGGCCATCTCAGATGCAGAG 508
Qy 361 GACTGCTGTCAGTGAACACGATACAGCTGAGATGGCAAGGGTGCCTACGGT 420
Db 509 GACTGCTGTCAGTGAACACGATACAGCTGAGATGGCAAGGGTGCCTACGGT 568
Qy 421 GTGTGTAGGCTGGCTTACAAAGTGAACAGACACTATGCAATGAAAGTCTCTTCC 480
Db 569 GTGTGTAGGCTGGCTTACAAAGTGAACAGACACTATGCAATGAAAGTCTCTTCC 628
Qy 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGGTCCAG 540
Db 629 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGGTCCAG 688
Qy 541 GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 600
Db 689 GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 748
Qy 601 GCCATCTGAAGAAGCTGGACCACTGAATGTGTCAAACTGATCGAGGTCTCTGATGAC 660
Db 749 GCCATCTGAAGAAGCTGGACCACTGAATGTGTCAAACTGATCGAGGTCTCTGATGAC 808
Qy 661 CCAGCTGAGGACAACTCTATTGTGTGTGACCTCTCGAGAAAGGGCCCGTCAATGAA 720
Db 809 CCAGCTGAGGACAACTCTATTGTGTGTGACCTCTCGAGAAAGGGCCCGTCAATGAA 868
Qy 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCCTCACTCGGGAGCTCATC 780
Db 869 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCCTCACTCGGGAGCTCATC 928
Qy 781 CTGGGCTTCGAGTACTTGCATCTGCAGAGATCGTCCACAGGACATCAAGCCATCCAAC 840
Db 929 CTGGGCTTCGAGTACTTGCATCTGCAGAGATCGTCCACAGGACATCAAGCCATCCAAC 988
Qy 841 CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGACTTTGGCGTCAGCAACCACTTT 900
Db 989 CTGCTCTCGGGGATGATGGGCACTGGAAGATCGCGACTTTGGCGTCAGCAACCACTTT 1048
Qy 901 GAGGGGAGGAGCTCAGCTGTCAGCAGCGGGGAAACCCAGCATTCATGGCCCCCGAG 960
Db 1049 GAGGGGAGGAGCTCAGCTGTCAGCAGCGGGGAAACCCAGCATTCATGGCCCCCGAG 1108
Qy 961 GCATTTCTGATTCGGGCGAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1020
Db 1109 GCATTTCTGATTCGGGCGAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1168
Qy 1021 GTACGTTGTACTGCTTTGTATGGGAAGTGCCTCATTCAGAGATTTGATCTCTGGCC 1080
Db 1169 GTACGTTGTACTGCTTTGTATGGGAAGTGCCTCATTCAGAGATTTGATCTCTGGCC 1228
Qy 1081 CTCCACAGGAAGATCAAGAAATGAGCCGCTGGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1229 CTCCACAGGAAGATCAAGAAATGAGCCGCTGGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1288
Qy 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTG 1200
Db 1289 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAAGATCCCGAGACGAGAAATGGGGTG 1348
Qy 1201 CCAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1260
Db 1349 CCAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1408
Qy 1261 GAGGAGCACTGAGCGTGTGAGGTGACAGAGGGGAGGTTTAAAGACTCAGTCAAGGCTC 1320
Db 1409 GAGGAGCACTGAGCGTGTGAGGTGACAGAGGGGAGGTTTAAAGACTCAGTCAAGGCTC 1468
Qy 1321 ATCCCCAGCTGACCAAGCTGATCTCTGTGTAAGTCCATGCTGAGGAAGGCTTCTTTGGG 1380
Db 1469 ATCCCCAGCTGACCAAGCTGATCTCTGTGTAAGTCCATGCTGAGGAAGGCTTCTTTGGG 1528
Qy 1381 AACCCGTTTGGCCCCCAGCGAGGAGGAGCGGATCCATCTCTGCTCCAGGAACCTTA 1440
Db 1529 AACCCGTTTGGCCCCCAGCGAGGAGGAGCGGATCCATCTCTGCTCCAGGAACCTTA 1588

Qy 1441 CTGCTGAAAGAGCGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGCTCCAGGAA 1500
Db 1589 CTGCTGAAAGAGCGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGCTCCAGGAA 1648
Qy 1501 GACGAGGCTGCATCC 1515
Db 1649 GACGAGGCTGCATCC 1663
RESULT 7
AX746188 1937 bp DNA linear PAT 13-JUN-2003
LOCUS AX746188
DEFINITION Sequence 39 from Patent WO0208399.
ACCESSION AX746188
VERSION AX746188.1 GI:31746174
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Yue, H., Khan, F.A., Gururajan, R., Hafalia, A.J., Chawla, N.K.,
Arvizu, C.S., Rankumar, J., Gandhi, A.R., Policky, J.L., Baughn, M.R.,
Tribouley, C.M., Bandman, O., Nguyen, D.B., Lu, Y., Burford, N., Lal, P.,
Ding, L., Yao, M.G., Elliott, V.S., Recipon, S.A., Kearney, L., Lu, D.A.,
Greenwald, S.R., Tang, Y.T., Xu, Y., Walsh, R.T., Gietzen, K.J.,
Yang, J., Jackson, J.L. and Thornton, M.
Human kinases
TITLE Patent: WO 0208399-A 39 31-JAN-2002;
JOURNAL Incyte Genomics, Inc. (US) ; Thornton, Michael (US)
FEATURES
source 1. 1937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 2190612CB1"

Query Match 99.7%; Score 1510.2; DB 6; Length 1937;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGAGGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGGTAGAAGCGGTG 60
Db 65 ATGAGGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGGTAGAAGCGGTG 124
Qy 61 GCAGCCATCGATGCTACTCTTGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 120
Db 125 GCAGCCATCGATGCTACTCTTGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 184
Qy 121 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGACCCCCACACCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 244
Qy 181 CTGCTCCAGGCCCGCTTAGCTCTCTCAGCAGGAGCTTCCCTACAGGAGCGGCAGCA 240
Db 245 CTGCTCCAGGCCCGCTTAGCTCTCTCAGCAGGAGCTTCCCTACAGGAGCGGCAGCA 304
Qy 241 GGAAGCTACTTGAGGCGCAGGCTGGGCTTATGCCAGGGGCTTCCAGCCACATCTCC 300
Db 305 GGAAGCTACTTGAGGCGCAGGCTGGGCTTATGCCAGGGGCTTCCAGCCACATCTCC 364
Qy 301 CCCGGGCTTGCGGAGGCCACCATCGAGTCCACACACGCTGGCCATCTCAGATCAGAG 360
Db 365 CCCGGGCTTGCGGAGGCCACCATCGAGTCCACACACGCTGGCCATCTCAGATCAGAG 424
Qy 361 GACTGCTGACGTGAACCACTGACAGCTGACAGTGAAGTGGCAAGGGTGCCTTACGGT 420
Db 425 GACTGCTGACGTGAACCACTGACAGCTGACAGTGAAGTGGCAAGGGTGCCTTACGGT 484
Qy 421 GTGCTGAGGCTGGCTTACAGGAAAGTGAAGACACACTATGCAATGAAAGTCTTTCC 480

Db 485 GTGGTGAAGGCTGGCCCTACAAACGAAAGTGAAGACAGACACTATCAATGAAGTCTCTTC 544

QY 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCCGAGAGGGTCCCCAG 540

Db 545 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCCGAGAGGGTCCCCAG 604

QY 541 GGTGCCACAGGAGGACAGCAAGCAGCTGTCTGCCCTGGAGCGGGTGTACCAAGAGATT 600

Db 605 GTGCCACAGGAGGACAGCAAGCAGCTGTCTGCCCTGGAGCGGGTGTACCAAGAGATT 664

QY 601 GCATCCTGAAGAGCTGGACACAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 660

Db 665 GCATCCTGAAGAGCTGGACACAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 724

QY 661 CCAGCTGAGGACAACTCTATTATTGGTGTGTGACCTCTGAGAAAGGGCGCGTCATGGAA 720

Db 725 CCAGCTGAGGACAACTCTATTATTGGTGTGTGACCTCTGAGAAAGGGCGCGTCATGGAA 784

QY 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCTCTAOCCTGCGGACGTCATC 780

Db 785 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCTCTAOCCTGCGGACGTCATC 844

QY 781 CTGGGCTCGAGTACTGACATGCCAGAGATCGTCCACAGGACATCAAGCCATCCAAC 840

Db 845 CTGGGCTCGAGTACTGACATGCCAGAGATCGTCCACAGGACATCAAGCCATCCAAC 904

QY 841 CTGCTCTGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 900

Db 905 CTGCTCTGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 964

QY 901 GAGGGAAACGACGCTCAGCTGTCCAGACGGCGGAAACCCAGCATTCATGAGCCCGGAG 960

Db 965 GAGGGAAACGACGCTCAGCTGTCCAGACGGCGGAAACCCAGCATTCATGAGCCCGGAG 1024

QY 961 GCATTTCTGATTTCGGCCAGAGCTTCACTGGGAGGCGCTTGATGTATGGGCCACTGGC 1020

Db 1025 GCATTTCTGATTTCGGCCAGAGCTTCACTGGGAGGCGCTTGATGTATGGGCCACTGGC 1084

QY 1021 GTCACTTTACTGCTTTGTCTATGGGAAGTGCCTATTCATGACGATTCATCTCTGGCC 1080

Db 1085 GTCACTTTACTGCTTTGTCTATGGGAAGTGCCTATTCATGACGATTCATCTCTGGCC 1144

QY 1081 CTCACAGAAATCAAGATGAGCCCGTGGTGTTCCTGAGAGCCAGAAATCAGCGAG 1140

Db 1145 CTCACAGAAATCAAGATGAGCCCGTGGTGTTCCTGAGAGCCAGAAATCAGCGAG 1204

QY 1141 GAGCTCAAGGACTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAAATTCGGGTG 1200

Db 1205 GAGCTCAAGGACTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAAATTCGGGTG 1264

QY 1201 CCAGACATCAAGTTGCACCTTTGGGTGACCAAGACGGGGAGGAGCCCTTCCTTCGGAG 1260

Db 1265 CCAGACATCAAGTTGCACCTTTGGGTGACCAAGACGGGGAGGAGCCCTTCCTTCGGAG 1324

QY 1261 GAGGAGCACTGACGGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

Db 1325 GAGGAGCACTGACGGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1384

QY 1321 ATCCCCAGCTGACCAACGCTGATCCTGGTGAAGTCCATCTGAGGAAGGTTCCCTTGG 1380

Db 1385 ATCCCCAGCTGACCAACGCTGATCCTGGTGAAGTCCATCTGAGGAAGGTTCCCTTGG 1444

QY 1381 AACCCGTTTGAAGCCCGACGAGGAGGAGCGGATCCATCTGCTCCAGGAACCTTA 1440

Db 1445 AACCCGTTTGAAGCCCGACGAGGAGGAGCGGATCCATCTGCTCCAGGAACCTTA 1504

QY 1441 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCTCCAGGAA 1500

Db 1505 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCTCCAGGAA 1564

QY 1501 GACGAGGCTGCATCC 1515

Db 1565 GACGAGGCTGCATCC 1579

RESULT 8
AX455763
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match
Best Local Similarity
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAACGGTG 60
Db 1 ATGGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAACGGTG 60
QY 61 GCAGCCATCGATGCTACTCTACTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAC 120
Db 61 GCAGCCATCGATGCTACTCTACTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAC 120
QY 121 GGTGTGGACCCCCACACGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 121 GGTGTGGACCCCCACACGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
QY 181 CTGCTCCAGACCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGAGCGGCAGCA 240
Db 181 CTGCTCCAGACCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGAGCGGCAGCA 240
QY 241 GGAAGCTATCTGGAGCGGCGAGCTGGGCTTATGCCACGGGGCCCTGCCAGCCATCTCC 300
Db 241 GGAAGCTATCTGGAGCGGCGAGCTGGGCTTATGCCACGGGGCCCTGCCAGCCATCTCC 300
QY 301 CCCCCGGGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db 301 CCCCCGGGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
QY 361 GACTCGGTGAGCTGAAACAGTCAAGCTGAGAGTGAGATTGGCAAGGGTGCCTACGGT 420
Db 361 GACTCGGTGAGCTGAAACAGTCAAGCTGAGAGTGAGATTGGCAAGGGTGCCTACGGT 420
QY 421 GTGGTGAAGCTGGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTCC 480
Db 421 GTGGTGAAGCTGGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTCC 480
QY 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCCGAGAGGGTCCCAG 540
Db 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCCGAGAGGGTCCCAG 540
QY 541 GGTGCCACAGGAGGACAGCAAGCAGCTGTCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
Db 541 GGTGCCACAGGAGGACAGCAAGCAGCTGTCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
QY 601 GCCATCTGAAGAGCTGGACCAAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 660
Db 601 GCCATCTGAAGAGCTGGACCAAGTGAATGTGGTCAAACTGATCAGAGTCTCTGGATGAC 660

QY 661 CCAGCTGAGGACAACTCTATTGGTGTGTTGACCTCTGAGAAAGGGCGCGTCATGGAA 720
Db 661 CCAGCTGAGGACAACTCTATTGGTGTGTTGACCTCTGAGAAAGGGCGCGTCATGGAA 720
QY 721 GTGCGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACCTCGGCGAGTCAATC 780
Db 721 GTGCGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACCTCGGCGAGTCAATC 780
QY 781 GTGCGCTGTGAGTCTGTCAGTCCAGCAAGATGCTCCACAGGAGCATCAAGCCATCCAAAC 840
Db 781 GTGCGCTGTGAGTCTGTCAGTCCAGCAAGATGCTCCACAGGAGCATCAAGCCATCCAAAC 840
QY 841 CTGCTCTCGGGGATGATGGGACGTGAAGATCGCGACTTTGGCGTCAGCAACACAGTTT 900
Db 841 CTGCTCTCGGGGATGATGGGACGTGAAGATCGCGACTTTGGCGTCAGCAACACAGTTT 900
QY 901 GAGGGGAACGAGCTGAGCTGTCAGCAGCAGCGCGGGAACCCAGCATTCATGGCCCCGAG 960
Db 901 GAGGGGAACGAGCTGAGCTGTCAGCAGCAGCGCGGGAACCCAGCATTCATGGCCCCGAG 960
QY 961 GCCATTTCTGATTCGGGCGAGAGCTTGTGCGGGAAGGCTTGGATGTATGGGCGACTGGC 1020
Db 961 GCCATTTCTGATTCGGGCGAGAGCTTGTGCGGGAAGGCTTGGATGTATGGGCGACTGGC 1020
QY 1021 GTCAAGTTGATCTGCTTGTCTATGCGGAAGTGCCTATTCAGTCCAGGATTTTCATCTGSCC 1080
Db 1021 GTCAAGTTGATCTGCTTGTCTATGCGGAAGTGCCTATTCAGTCCAGGATTTTCATCTGSCC 1080
QY 1081 CTCCACAGGAGATCAAGAATGAGCCGTGTGTCTTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1081 CTCCACAGGAGATCAAGAATGAGCCGTGTGTCTTCTGAGGAGCCAGAAATCAGCGAG 1140
QY 1141 GAGCTCAGGACCTGATCTCAAGATGTTAGCAAGATTCGAGACGAGAAATGGGGTG 1200
Db 1141 GAGCTCAGGACCTGATCTCAAGATGTTAGCAAGATTCGAGACGAGAAATGGGGTG 1200
QY 1201 CCAGACATCAAGTTGCACTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCCGGAG 1260
Db 1201 CCAGACATCAAGTTGCACTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCCGGAG 1260
QY 1261 GAGGAGCACTGACGCTGTTGGAGGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1261 GAGGAGCACTGACGCTGTTGGAGGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
QY 1321 ATCCCGAGCTGACGAGTATCTGTTGAAGTCCATGCTGAGGAGCGTTCCTTTGGG 1380
Db 1321 ATCCCGAGCTGACGAGTATCTGTTGAAGTCCATGCTGAGGAGCGTTCCTTTGGG 1380
QY 1381 AACCCGTTTGAGCCCGGACGAGGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1440
Db 1381 AACCCGTTTGAGCCCGGACGAGGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1440
QY 1441 CTGTTGAAGAGGGTTGTTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCCAGGAA 1500
Db 1441 CTGTTGAAGAGGGTTGTTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCCAGGCT 1500
QY 1501 GAGCA 1505
Db 1501 TACCA 1505

RESULT 9
AX455761
LOCUS AX455761 2711 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 7 from Patent WO0220800.
ACCESSION AX455761
VERSION AX455761.1 GI:21714788
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1

AUTHORS Meyers, R.A. and Silos-Santiago, I.
TITLE 1658, 14223, and 16002, novel human kinases and uses therefor
JOURNAL Patent: WO 0220800-A 7 14-MAR-2002;
MILLENNIUM PHARM INC (US)
FEATURES
source 1. 2711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
198. .1883
/note="unnamed protein product"
Location/Qualifiers
1. 2711
/codon_start=1
/protein_id="CAD38244.1"
/db_xref="GI:21714789"
/db_xref="REMBL:CAD38244"
/translation="MEGGPACVCCDPAELVVERVAADIVTHLEADGGPPTRNQVDP
PPRAASVIPGSTSRLLPAPSLARKLSLOERPAGSYLEAQAPYATGPASHISPR
APRRPTTSEHVATISDAEDCVLNQYKLSIIEIGKAYGVVRLAYNESDRHYAMKVL
KKLLKQYGFPRPSPRSGOAGGPAKQLPLERVOEAILKKLDHNVVKLIEVL
DDPAEDNLVLFDFLLRKGPVMEPCDKPFSEQARLYLRDVLGLEVLCQKIVHRDI
KPSNLLGDDGHVKIADFGVSNOPEDNAQLSSTAGTAPMAPEAISDSGQSFSGKAL
DVWATGVLLYCFVYKCPFIDDFILALHRKIKNEPVVPFEGPEISEKLDLILKWLK
NPETRIGVDPDIKLHPWTNKEEPLSEBEHCNVVEEVEKNLSVLIFSWMTVILV
KSMRKESFNGNPFPPQARBERSMSAPENLIVKEFGEGKSPGLPGVQAVYHNRDLP
LGPRPGSAATGIWRGYPFAHLGVVAPELLSIQTLVWGLGLTGDRAVSVEF"
ORIGIN
Query Match 98.4%; Score 1490.6; DB 6; Length 2711;
Best Local Similarity 99.4%; Pred. No. 9.1e-312;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGCGAGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
Db 198 ATGAGGGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 257
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGCTGCCCCAGAGCCCTACTAGAAAC 120
Db 258 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGCTGCCCCAGAGCCCTACTAGAAAC 317
QY 121 GGTGTGACACCCCAACACGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 318 GGTGTGACACCCCAACACGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 377
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGAGCTTCCCTACAGGAGCGCCAGCA 240
Db 378 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGAGCTTCCCTACAGGAGCGCCAGCA 437
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTTGCACGCCATCTCC 300
Db 438 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTTGCACGCCATCTCC 497
QY 301 CCCGGGCTGGCGGAGGCCACCATCGAGTCCCAACCCAGCTGGCCATCTCAGATGCGAG 360
Db 498 CCCGGGCTGGCGGAGGCCACCATCGAGTCCCAACCCAGCTGGCCATCTCAGATGCGAG 557
QY 361 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACCGT 420
Db 558 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACCGT 617
QY 421 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACATATATGCAATGAAGTCTTCTCC 480
Db 618 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACATATATGCAATGAAGTCTTCTCC 677
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCCTCCCGGAGGGTCCCGAG 540
Db 678 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCCTCCCGGAGGGTCCCGAG 737
QY 541 GCTGCCAGGAGGAGGAGCCAGCAAGCAGTGTGCTCCCTGGAGCGGGTGTACAGAGATT 600
Db 738 GCTGCCAGGAGGAGGAGGAGCCAGCAAGCAGCTGTGCCCTGGAGCGGGTGTACAGAGATT 797
QY 601 GCCATCTCTGAAGAGCTGGACCACTGAATGTGTGCTCAACTGATCGAGGTCTCTGATGAC 660

QY	814	GTCCACAGGGGACATCAAGCCATCCAACTGCTCCTGGGGATGATGGCCACGTTGAAGATC	873
Db	838	GTCCACAGGGGACATCAAGCCATCCAACTGCTCCTGGGGATGATGGCCACGTTGAAGATC	897
QY	874	GCCGACTTTGGCGTCAGCAACACAGTTTGAGGGGAACAGCGTCAAGCTGTCCAGCACGGCG	933
Db	898	GCCGACTTTGGCGTCAGCAACACAGTTTGAGGGGAACAGCGTCAAGCTGTCCAGCACGGCG	957
QY	934	GGAAACCCAGCAATTCATGGCCCCCGAGGCCAATTTCTGATTCGGGCCAGAGCTTCAGTGGG	993
Db	958	GGAAACCCAGCAATTCATGGCCCCCGAGGCCAATTTCTGATTCGGGCCAGAGCTTCAGTGGG	1017
QY	994	AAGGCTTGGATGATGGGCCACTGGCGTCAAGTGTACTGCTTTGTCTATGGGAAGTGC	1053
Db	1018	AAGGCTTGGATGATGGGCCACTGGCGTCAAGTGTACTGCTTTGTCTATGGGAAGTGC	1077
QY	1054	CCATTTCATCGACGATTTTCATCCTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG	1113
Db	1078	CCATTTCATCGACGATTTTCATCCTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG	1137
QY	1114	TTTCTGAGGAGCCAGAAATCAGCGAGAGCTCAAGGACCTGATCCTGAAGATGTTAGAC	1173
Db	1138	TTTCTGAGGAGCCAGAAATCAGCGAGAGCTCAAGGACCTGATCCTGAAGATGTTAGAC	1197
QY	1174	AAGANTCCGAGACGAGAAATGGGGTGCAGACATCAAGTTGCACTTGGGTGACCAAG	1233
Db	1198	AAGANTCCGAGACGAGAAATGGGGTGCAGACATCAAGTTGCACTTGGGTGACCAAG	1257
QY	1234	AACGGGGAGGAGCCCTTCCTTCGGAGGAGGAGCACTGCAGCGTGGTGAGGAGTGCAGAG	1293
Db	1258	AACGGGGAGGAGCCCTTCCTTCGGAGGAGGAGCACTGCAGCGTGGTGAGGAGTGCAGAG	1317
QY	1294	GGGAGGTTAAGAACTCAGTAGGCTCATPCCCACTGGAACCAACCGTGTATCTGTGTGAAG	1353
Db	1318	GAGGAGTTAAGAACTCAGTAGGCTCATPCCCACTGGAACCAACCGTGTATCTGTGTGAAG	1377
QY	1354	TCCATGCTGAGGAGCGTTCTTTTGGGAACCCGTTTGAGCCCGGACGAGGAGGAAGAG	1413
Db	1378	TCCATGCTGAGGAGCGTTCTTTTGGGAACCCGTTTGAGCCCGGACGAGGAGGAAGAG	1437
QY	1414	CGATCCATGCTCTGCTCCAGGAACCTACTGTGTGAAGAGGGTTTGTGAAGGGGGCAAG	1473
Db	1438	CGATCCATGCTCTGCTCCAGGAACCTACTGTGTGAAGAGGGTTTGTGAAGGGGGCAAG	1497
QY	1474	AGCCAGAGCTCCCGGGCTCCAGGAAGCAGAGGCTGCATCC	1515
Db	1498	AGCCAGAGCTCCCGGGCTCCAGGAAGCAGAGGCTGCATCC	1539
RESULT 11			
AXI179641			
LOCUS			
Sequence 14 from Patent WO0146397.			
AXI179641			
ACCESSION			
AXI179641.1			
VERSION			
GI:15132054			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Yang, J., Raughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Reddy, R.,			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Yue, H., Yao, M.G., Lal, P. and Khan, F.A.			
Human Kinases			
Patent: WO 0146397-A 14 28-JUN-2001;			
Incyte Genomics, Inc. (US)			
FEATURES			
Location/Qualifiers			
1..3501			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="Incyte ID No: 2041716CB1"			
source			
ORIGIN			
Query Match 96.2%; Score 1457.2; DB 6; Length 3501;			
Best Local Similarity 97.9%; Pred. No. 1.5e-304; Indels 30; Gaps 2;			
Matches 1509; Conservative 0; Mismatches 3			
QY	1	ATGAGAGGGGTCAGAGTGTCTGTCAGAGATCCTCGGGCAGAGCTGGTGAAGACGGGTG	60
Db	65	ATGAGAGGGGTCAGAGTGTCTGTCAGAGATCCTCGGGCAGAGCTGGTGAAGACGGGTG	124
QY	61	GCAGCCATCGATGATGATCTACTTGGAGAGGACATGGTGGCCAGAGCCCTACTAGAAC	120
Db	125	GCAGCCATCGATGATGATCTACTTGGAGAGGACATGGTGGCCAGAGCCCTACTAGAAC	184
QY	121	GCTGTGAGACCCCCACACAGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	180
Db	185	GCTGTGAGACCCCCACACAGGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	244
QY	181	CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	240
Db	245	CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	304
QY	241	GGAGACTATCTGAGAGCGCAGAGCTGGGCTTTATGCCAGGGGCTGCCAGCCACATCTCC	300
Db	305	GGAGACTATCTGAGAGCGCAGAGCTGGGCTTTATGCCAGGGGCTGCCAGCCACATCTCC	364
QY	301	CCCGGGCTCGCGAGGGCCACCATCGAGTCCCACACACGTGGCCATCTCAGATGACAGAG	360
Db	365	CCCGGGCTCGCGAGGGCCACCATCGAGTCCCACACACGTGGCCATCTCAGATGACAGAG	424
QY	361	GACTGCGTGCAGCTGAACCCAGTACAAGCTGAGAGTGAGATTGGCA-----	406
Db	425	GACTGCGTGCAGCTGAACCCAGTACAAGCTGAGAGTGAGATTGGCAAGGTGGGGCTGACT	484
QY	407	-----AGGGTGCCTACGGTGTGCTGAGGCTGGCCTACAACGAAGTGAAGAC	453
Db	485	GATGCTATCTGACAGGTCCTACGGTGTGAGGCTGGCCTACAACGAAGTGAAGAC	544
QY	454	AGACACTATGCAATGAAGTCTTTCCAAAAGAGTTACTGAAGCAGTATGGCTTCCA	513
Db	545	AGACACTATGCAATGAAGTCTTTCCAAAAGAGTTACTGAAGCAGTATGGCTTCCA	604
QY	514	CGTGGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACACAGCAGCAGCTGCTG	573
Db	605	CGTGGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACACAGCAGCAGCTGCTG	664
QY	574	CCCTGGAGCGGCTTACAGAGATTCCTGTAAGAGCTGCAACACCTGATGATG	633
Db	665	CCCTGGAGCGGGTGTACAGAGATTCCTGTAAGAGCTGCAACACCTGATGATG	724
QY	634	GTCAAACTGATCGAGGTCCTGGATGACCCAGCTGAGGACAACTCTATTGGTGTGAC	693
Db	725	GTCAAACTGATCGAGGTCCTGGATGACCCAGCTGAGGACAACTCTATTGGTGTGAC	781
QY	694	CTCCTGAGAAAGGGCCCGTGTATGAAAGTCCCTGTGCAAGCCCTTCTCGAGAGGCA	753
Db	782	CTCCTGAGAAAGGGCCCGTGTATGAAAGTCCCTGTGCAAGCCCTTCTCGAGAGGCA	841
QY	754	GCTCGCCTCTACTCGCGGACGTCTCTGGGCTCTGAGTACTTGCACCTCCAGAGATC	813
Db	842	GCTCGCCTCTACTCGCGGACGTCTCTGGGCTCTGAGTACTTGCACCTCCAGAGATC	901
QY	814	GTCCACAGGACATCAAGCCATCCAACTGCTCTGGGGATGATGGGCACTGTAAGATC	873
Db	902	GTCCACAGGACATCAAGCCATCCAACTGCTCTGGGGATGATGGGCACTGTAAGATC	961
QY	874	GCCGACTTTGGCGTCAGCAACACAGTTTGAGGGGAACAGCGTCAAGTGTCCAGCAGCGG	933
Db	962	GCCGACTTTGGCGTCAGCAACACAGTTTGAGGGGAACAGCGTCAAGTGTCCAGCAGCGG	1021
QY	934	GGAAACCCAGCAATTCATGGGCCCGGAGGCCATTTCTGATTCGGGCCAGAGCTTCAAGTGG	993
Db	1022	GGAAACCCAGCAATTCATGGGCCCGGAGGCCATTTCTGATTCGGGCCAGAGCTTCAAGTGG	1081

QY	994	AGGCCCTTGATATATGGCCACTGGCGTCAACGTTGTACTGCTTTGTCTATGGAAAGTGC	1053
Db	1082	AGGCCCTTGATATATGGCCACTGGCGTCAACGTTGTACTGCTTTGTCTATGGAAAGTGC	1141
QY	1054	CGATTTCATGACGATTTTCATCTCTGGCCCTCCAGGAAGATCAAGATGAGCCCGTGGT	1113
Db	1142	CGTTTCATGACGATTTTCATCTCTGGCCCTCCAGGAAGATCAAGATGAGCCCGTGGT	1201
QY	1114	TTTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC	1173
Db	1202	TTTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC	1261
QY	1174	AGAATCCCGAGACGAGAAATGGGGTCCAGACATCAAGTTGACCCCTTGGGTGACCAAG	1233
Db	1262	AGAATCCCGAGACGAGAAATGGGGTCCAGACATCAAGTTGACCCCTTGGGTGACCAAG	1321
QY	1234	ACGGGGAGGAGCCCTTCTCTCGGAGGAGGACACTGACGCTGGTGGAGGTGACAGAG	1293
Db	1322	ACGGGGAGGAGCCCTTCTCTCGGAGGAGGACACTGACGCTGGTGGAGGTGACAGAG	1381
QY	1294	GGGGAGGTTAAGAACTCAGTCAAGCTCATCCCGAGCTGGACACGCTGATCTCTGGTGAAG	1353
Db	1382	GAGGAGTTAAGAACTCAGTCAAGCTCATCCCGAGCTGGACACGCTGATCTCTGGTGAAG	1441
QY	1354	TCATGCTGAGGAGGCTTCTCTTGGAAACCCGTTTGGCCCGGACGAGGAGGAGAG	1413
Db	1442	TCATGCTGAGGAGGCTTCTCTTGGAAACCCGTTTGGCCCGGACGAGGAGGAGAG	1501
QY	1414	CGATCATCTCTGCTCCAGGAACCTTCTGGTGAAGAGGGTTTGGTGAAGGGGCAAG	1473
Db	1502	CGATCATCTCTGCTCCAGGAACCTTCTGGTGAAGAGGGTTTGGTGAAGGGGCAAG	1561
QY	1474	AGCCGAGAGCTCCCGGCGCTCAGGAAGACGAGGCTGCATCC	1515
Db	1562	AGCCGAGAGCTCCCGGCGCTCAGGAAGACGAGGCTGCATCC	1603
RESULT 12			
LOCUS	BC031647		
DEFINITION	Homo sapiens calcium/calmodulin-dependent protein kinase kinase 1, alpha, transcript variant 3, mRNA (cDNA clone MGC:34095 IMAGE:5175128), complete cds.		
ACCESSION	BC031647		
VERSION	BC031647.1	GI:21594835	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2469)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Ditzchenko, L., Mariani, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Lequellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalski, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

22388257
12477932
2 (bases 1 to 2469)
Strausberg, R.

Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lousseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAK Plate: 51 Row: 1 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27437010.

FEATURES
source

Location/Qualifiers

1..2469

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/clone="MGC:34095 IMAGE:5175128"

/tissue_type="Brain, lung, Testis, adult, pooled whole"

/clone_lib="NIH MGC_115"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1..2469

/gene="CAMKK1"

/note="synonyms: DKFp761M0423, CAMKKA, MGC34095"

/db_xref="LocusID:84254"

83..1645

/codon_start=1

/product="calcium/calmodulin-dependent protein kinase 1 alpha, isoform b"

/protein_id="AAH31647.1"

/db_xref="GI:21594836"

/db_xref="LocusID:84254"

/translation="MEGGPVCDDPRAELVERVAALDVTHLEADGGPEPRNGVDP

PPRAASVPGSTSRLLPAPLSARKLSQERPAQYATGPAHSLSPR

AWRPTTSHHVAISDAEDCVLNQYKLSGKAGVVRVLAAYNESDRHYAKVL

KKLKLKQYFPRPPRPSQAQGGPAKQLLERYQEIATLKKLHVNVVKLEVL

DDPADNLYLALQNAQIQLDSTNIAKPHSLPSEQDGSSTWAARSVDFLLKRGV

MEVCDKPFSEQARLYRDLVLELHCKIVHRDIKPSNLLGDGDHVKIADFGV

SNQPCDQALSTAGTAFNAWPAISDGSQFSKGLDVATGVTLYCFVYKGPFI

DDFILAHRKIKNPVVPFPEPELSEELKDLILKMLDKNPETRGVDPDKLHPWTKN

ERSPLPSEEEHCSVVEVEEVKNSVRLIPSWTTVILVKSMLRKRSGNPFPEQARRE

464..1423

/note="s TKc; Region: Serine/Threonine protein kinases, catalytic domain"

/db_xref="CDD:smart00220"

ORIGIN

Query Match 87.2%; Score 1321.2; DB 9; Length 2469;

Best Local Similarity 92.0%; Pred. No. 3.9e-275;

Matches 1453; Conservative 0; Mismatches 13; Indels 114; Gaps 1;

QY 1 ATGGAGGGGGTCCAGCTGTCTGTGCGAGATCTCGGGCAGAGCTGTGTAAGACGGGTG 60

Db 83 ATGGAGGGGGTCCAGCTGTCTGTGCGAGATCTCGGGCAGAGCTGTGTAAGACGGGTG 142

misc_feature

QY 61 GCAGCCATCGATGCTGACTCACTTGGAGGAGCGCAGATGCTGGCCCGCAGAGCCCTACTAGAAC 120
Db 143 GCAGCCATCGATGCTGACTCACTTGGAGGAGCAGATGCTGGCCCGCAGAGCCCTACTAGAAC 202
QY 121 GGTGTGAGACCCCCACACAGCGGCCAGAGCTGCCTCTGTGATCCCTTGGCAGTACTTCAAGA 180
Db 203 GGTGTGNACCCCCACACAGCGGCCAGAGCTGCCTCTGTGATCCCTTGGCAGTACTTCAAGA 262
QY 181 CTGCTCCAGCCCGCCCTAGCTCTTCAAGCAGGAGAGCTTTCCTTACAGGAGCGGCCAGCA 240
Db 263 CTGCTCCAGCCCGCCCTAGCTCTTCAAGCAGGAGAGCTTTCCTTACAGGAGCGGCCAGCA 322
QY 241 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGCGACGGGSCCTGCCAGCCACATCTCC 300
Db 323 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGCGACGGGSCCTGCCAGCCACATCTCC 382
QY 301 CCCCGGGCTTGGCGAGGCCACCATCGAGTCCCAACACGTCGGCCATCTCAGATGCGAG 360
Db 383 CCCCGGGCTTGGCGAGGCCACCATCGAGTCCCAACACGTCGGCCATCTCAGATGCGAG 442
QY 361 GACTGCGTGAGCTGAACAGTACAGCTGAGAGTGGCAAGGTCCTACGGT 420
Db 443 GACTGCGTGAGCTGAACAGTACAGCTGAGAGTGGCAAGGTCCTACGGT 502
QY 421 GTGTGTAGGCTGGCTTACAAAGAAAGTGAAGACACACTATGCAATGAAAGTCTTTC 480
Db 503 GTGTGTAGGCTGGCTTACAAAGAAAGTGAAGACACACTATGCAATGAAAGTCTTTC 562
QY 481 AAAAAGAAAGTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCGCAGAGGGTCCAG 540
Db 563 AAAAAGAAAGTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCGCAGAGGGTCCAG 622
QY 541 GCTGCCAGGAGGACGAGCCAGCAGCTGCTGCCCTGGAGCGGCTGACAGGAGATT 600
Db 623 GCTGCCAGGAGGACGAGCCAGCAGCTGCTGCCCTGGAGCGGCTGACAGGAGATT 682
QY 601 GCCATCTCTGAAGAAGCTGACACCTGAATGTGGTCAAACTGATCGAGGTCCTTGATGAC 660
Db 683 GCCATCTCTGAAGAAGCTGACACCTGAATGTGGTCAAACTGATCGAGGTCCTTGATGAC 742
QY 661 CCAGCTGAGGACAACTCTATTG----- 684
Db 743 CCAGCTGAGGACAACTCTATTGCGCCTCTGAGAACAGGCCACAGATATCCAGTTAGAT 802
QY 685 ----- 684
Db 803 TCACAAATATGCCAAGCCCACTCCCTGCTTCCTCTGAGCAGCAAGATGGATCC 862
QY 685 -----GTGTTTGACCTCTGAGAAAGGGGCCCGTTCATGGAAGTGCCC 726
Db 863 ACGTGGGCTGCGGCTCAGTTGACCTCTGAGAAAGGGGCCCGTTCATGGAAGTGCCC 922
QY 727 TGTGCAAGCCCTTCTCGGAGAGCAAGCTCGCTCTACCTGCGGGAAGTATCTCTGGG 786
Db 923 TGTGCAAGCCCTTCTCGGAGAGCAAGCTCGCTCTACCTGCGGGAAGTATCTCTGGG 982
QY 787 CTCGAGTACTTGCATGCGCAGAGATCGTCCACAGGGACATCAAGCCATCCAACTGCTC 846
Db 983 CTCGAGTACTTGCATGCGCAGAGATCGTCCACAGGGACATCAAGCCATCCAACTGCTC 1042
QY 847 CTGGGGGATGATGGGCACTGAAGATCGCCAGCTTGTGGCGTCAGCAACAGTTGAGGGG 906
Db 1043 CTGGGGGATGATGGGCACTGAAGATCGCCAGCTTGTGGCGTCAGCAACAGTTGAGGGG 1102
QY 907 AACGAGCTCAGCTTCAGCAGCGCGGGAACCCAGCATTCATGCCGCCCGAGCCATT 966
Db 1103 AACGAGCTCAGCTTCAGCAGCGCGGGAACCCAGCATTCATGGCCCCCGAGGCCATT 1162
QY 967 TCTGATTCGGGCCAGAGCTTCAGTGGGAGGCTTGGATGTATGGGCCACTGGCGTCAG 1026
Db 1163 TCTGATTCGGGCCAGAGCTTCAGTGGGAGGCTTGGATGTATGGGCCACTGGCGTCAG 1222

QY 1027 TTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCATCTCGCCCTCCAC 1086
Db 1223 TTGTACTGCTTTGTCTATGGGAAGTGCCCGTTTCATCGACGATTTTCATCTCGCCCTCCAC 1282
QY 1087 AGGAAGATCAAGAATAGCCCGTGTGTTTCTTCTGAGGAGCCAGAAATCAGCGAGGAGCTC 1146
Db 1283 AGGAAGATCAAGAATAGCCCGTGTGTTTCTTCTGAGGAGCCAGAAATCAGCGAGGAGCTC 1342
QY 1147 AAGGACCTGATCTTGAAGATGTTAGACAAGNAATCCGAGAGAGNAATGGGGTGCACAC 1206
Db 1343 AAGGACCTGATCTTGAAGATGTTAGACAAGNAATCCGAGAGAGNAATGGGGTGCACAC 1402
QY 1207 ATCAAGTTGCAACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTTCTCGAGGAGGAG 1266
Db 1403 ATCAAGTTGCAACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTTCTCGAGGAGGAG 1462
QY 1267 CACTGCAAGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1326
Db 1463 CACTGCAAGCTGTGGAGGTGACAGAGGAGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1522
QY 1327 AGCTGGACCCAGGTCCTGTCGAGTCCATGCTGAGGAAGCGTTCCTTTGGGAACCCG 1386
Db 1523 AGCTGGACCCAGGTCCTGTCGAGTCCATGCTGAGGAAGCGTTCCTTTGGGAACCCG 1582
QY 1387 TTTGAGCCCCCAGGCAAGGAGGAGGATGCCATGTCCTCCAGGAAACCTACTGTG 1446
Db 1583 TTTGAGCCCCCAGGCAAGGAGGAGGATGCCATGTCCTCCAGGAAACCTACTGTG 1642
QY 1447 AAAGAAGGTTTGGTGAAG 1466
Db 1643 TAAGTACTGTGGCGCAGG 1662
RESULT 13
BC017529
LOCUS
DEFINITION Mus musculus calcium/calmodulin-dependent protein kinase kinase 1,
alpha, mRNA (cdna clone MGC:27706 IMAGE:4924656), complete cds.
ACCESSION BC017529
VERSION BC017529.1 GI:17028423
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3435)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,F., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Wolfe,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Casavant,T.L.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smaluk,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
TITLE 2 (bases 1 to 3435)
JOURNAL Strausberg,R.
MEDLINE Direct Submission
REFERENCE 2
AUTHORS Submitted (19-NOV-2001) National Institutes of Health, Mammalian
JOURNAL

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Kryzanski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 38 Row: p Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

Location/Qualifiers

1. 3435

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:27706 IMAGE:4924656"

/tissue_type="Salivary gland, 10 week old female mouse"

/clone_lib="NCI CGAP_SG2"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 3435

/gene="Camk1"

/notes="synonyms: Camk1, CamKalpha"

/db_xref="LocusID:55984"

/db_xref="MGI:1891766"

127_1644

/codon_start=1

/product="Camk1 protein"

/protein_id="AAH17529.1"

/db_xref="GI:17028424"

/db_xref="LocusID:55984"

/translation="MESGAVCCQDPRAELVDRVAALNVHALEADEGPEPARNGVDP
PPRAASVIFGSAKRPTVPSLSARKFSLOERPAGSLGQAQVGYSTGPASHISPR
SWRPETIESHRVAISDEPCVQLNQKLSQEIKGAYGVRLAYNESEDRHYAMKVLIS
KKKLLKQYFRRPFRPGSQATQGGPAKQLPLERVYQRIAILKKLDHVNVLKLEVL
DDPAEDNLYLFDLLRKGPMVEVPCDPPPEQARLYRDIIILGLEYLHCOKIHRDI
KPSNLLLDGDDHVCVKADFGVSNQFRENDAQLSTAGTPAPMAPEAISDSGQSFSGAL
DWTATGVTLVCFVQKCPIDYIILTLHRKIKNEAVFPEPEVSEDLKDLIRMLDK
NPETRIGSDIKLHPWTKHGEPLFSEEEHCVSEVEEKNVSLPITVILV
KSLMKRSGNFPEFPOAREERSMSAPGSLMKEGCGCKSPFLPGVQDEAAS"

508. 1353

/note="S_TKc; Region: Serine/Threonine protein kinases, catalytic domain"

/db_xref="CDD:smart00220"

ORIGIN

Query Match 81.9%; Score 1241.4; DB 10; Length 3435;
Best Local Similarity 88.7%; Pred. No. 6.7e-258;
Matches 1344; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGCTCTGTCGCCAGGATCTCTCGGCGAGAGCTGTGTAGACGGGTG 60
DB 127 ATGGAGAGTGGCCCGCCGCTGCTGCCAAGACCTCGTGCAGAGCTGTGTGATCGGGTG 186
QY 61 GCAGCCATCGATGTGACTTCACTTGGAGGGGAGATGGTGGCCCGAGAGCTACTAGAAAC 120

REMARK COMMENT

187 GCGCCATCAATGTGGCCACTTGGAAAGCAGATGAGGGTCCAGAGCCTGCCAGGAAT 246
QY 121 GGTGTGAGCCCCCACCACACGGGGCAGAGCTCCTCTGTGTGATCCCTTGGCAGTACTTCAAGA 180
DB 247 GGTGTGATCTCCACCCCGGGCCAGAGCTGCTCGGTGATCCCTTGGCAGTCTTCAAGA 306
QY 181 CTGCTCCAGCCCGGCTAGCTCTCAGCAGAGGAGCTTTCCTTACAGAGGGGGCAGCA 240
DB 307 CCCACTCAGTGCAGCTTCCGCTTGAAGAGTTCCTCCCTGCAGAAACGACGCT 366
QY 241 GGAAGCTATCTGGAGGGCAGGCTTGGGCTTATGCAACGGGGCTGCCAGACCAATCTCC 300
DB 367 GGAAGCTGTCTGGGGCTCAAGTTGGGCTTACTCTACAGAGCTTCCAGTCACTCTCT 426
QY 301 CCCGGGCTTGGCGGAGGCCACCATCGAGTCCACCAACGCTGGCCATCTCAGATGCGAG 360
DB 427 CCTCGGTCTTGGCGGAGACCCACCATCGAGTCCCAACCGTGTGGCCATCTCAGACACAG 486
QY 361 GACTGCTGAGCTGAACCACTACAGCTCAAGCTGCAGAGTGCAGATTGCAAGGGTGCCTACGT 420
DB 487 GACTGTGTCAACTGAACCACTACAGCTGCAGAGTGCAGATTGCAAGGGTGCCTATGGT 546
QY 421 GTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGAGTCTTTCC 480
DB 547 GTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGAGTCTTTTC 606
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCACGTGGCCTCCCCCGAGAGGGTCCCGAG 540
DB 607 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCTCGCGCTCCTCCCCCAGAGGGTCCCAA 666
QY 541 GCTGCCCAGGGAGCAGCCACGACGCTGCTGCCCTTGGAGGGGTGTACAGAGGATT 600
DB 667 GCTACCCAGGGAGGGCCAGCCAAACAGCTGCTGCCCTTGGAGGGTGTGTATCAGGAGATT 726
QY 601 GCCATCTGAAGAAGCTGGACCAACGTAATGTGTGCTCAAACTGATCGAGGTCCTGGATGAC 660
DB 727 GCCATTCTAAAGAGCTGGACCAACGTAATGTGTGATGATGATGATGATGATGATGATGAT 786
QY 661 CCAGCTGAGACAACTCTATTGGTGTGTGACCTCTCAGAAAGGGGGCCGCTCATGGAA 720
DB 787 CCGCTGAAAGCAACCTCTATTGGTGTTCGACCTCTCAGAAAGGGGGCCGCTCATGGAA 846
QY 721 GTGCCCTGTGACAAGCCCTTCTCGAGAGGAGCAGCTCGCTCTACCTGGGGAGCTCATC 780
DB 847 GTGCCCTGCGACAGCCCTTCCAGAGGAGCAGCTCGCTCTACCTGGGGAGCTCATC 906
QY 781 CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTCCACAGGGACATCAAGCCATCAAC 840
DB 907 CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTCCACAGAGATGCTCCAGGGACATCAAGCCATCAAC 966
QY 841 CTGCTCTGGGGATGATGGGACGCTGAAGATCGCCGACTTTGGGCTCAGCAACCACTTT 900
DB 967 CTACTCTTGGGGACGAGCGGCATGTGAAGATCGCCGACTTTGGTGTCTAGTAACCACTTT 1026
QY 901 GAGGGAAACGACGCTCAGCTGTCCAGCAGCGGGGAAACCCAGCATTATATGCCCCCGAG 960
DB 1027 GAGGGAAATGATCTCAGCTGTCCAGTACGCGAGGAGCCCCGGCATTATGCCCCCGAG 1086
QY 961 GCCATTCTGATTCGGCCAGAGCTTCACTGGAAAGGCTTGGATGTATGGCCCACTGAC 1020
DB 1087 GCCATTTCTGATTCGGCCAGAGCTTCACTGGAAAGGCTTGGATGTATGGCCCACTGAC 1146
QY 1021 GTCAGTTGTACTGCTTGTCTATGGGAAGTGCCTCATTCAGCAGATTTTCATCTCTGGCC 1080
DB 1147 GTCAGCTATATGCTTGTCTATGGGAAGTGCCTCATTCAGTACTTGTACTTCTGAGG 1206
QY 1081 CTCACAGGAAGATCAAGAATGAGCCGCTGCTTCTCTGAGGAGCCAGAAATCAGCGAG 1140
DB 1207 CTGACCGTAAGATCAAGAATGAGCCGCTGCTTCTCTGAGGAGCCAGAGGTCAGCGAG 1266
QY 1141 GAGCTCAAGGACCTGATCCTCGAAGATGTAGCAAGAAATCCGAGACGAGAAATGGGGTG 1200

Db 1267 GACCTCAAGACCTGATCTCTGAGGATGCTAGACAAGATCTCTGACACAAGAAATGGGGTG 1326

Qy 1201 CCAGACATCAAGTTGACCTTGGGTGACCAAGAACCGGGAGGAGCCCTTCCCTTCGGAG 1260

Db 1327 TCAGATATCAAGTTACACCTTGGGTGACCAAGCATGGAGAGGAGCCCTTCCCTTCAGAG 1386

Qy 1261 GAGGAGCACTGAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

Db 1387 GAGGAGCACTGAGGTGGTGGAGGTGACCGAGGAGGAGGTGAAGAACTCAGTCAGGCTC 1446

Qy 1321 ATCCCACTGGACCAACGCTGATCTCTGAGTCCATCTGAGGAAGCTTCCCTTTGGG 1380

Db 1447 ATCCCACTGGACCAACGCTGATCTCTGAGTCCATCTGAGGAAGCTTCCCTTTGGG 1506

Qy 1381 AACCCGTTTGAAGCCCCAGGACCGGAGGAGGAGCGATCCATGTCCTGAGGAAGAACTTA 1440

Db 1507 AACCCATTCGAGCCCCAAGCAGCGAGGAGGAGGATCCATGTCCTGACCAAGGAGCTTA 1566

Qy 1441 CTGTGAAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1500

Db 1567 CTGATGAAAGAGGATGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1626

Qy 1501 GACGAGGCTGCATCC 1515

Db 1627 GATGAGGCTGCCTCC 1641

RESULT 14

LOCUS S83194

DEFINITION Ca2+/calmodulin-dependent protein kinase IV kinase isoform [rats,

ACCESSION S83194

VERSION S83194.1 GI:1836160

KEYWORDS Rattus sp.

SOURCE Rattus sp.

ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 3429)

AUTHORS Okuno,S., Kitani,T. and Fujisawa,H.

TITLE Evidence for the existence of Ca2+/calmodulin-dependent protein kinase IV kinase isoforms in rat brain

JOURNAL J. Biochem. 119 (6), 1176-1181 (1996)

MEDLINE 96425004

PUBMED 8827455

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 179633] from the original journal article.

FEATURES

source Location/Qualifiers

1..3429

/organism="Rattus sp."

/mol_type="mRNA"

/db_xref="taxon:10118"

1..3429

/gene="Ca2+/calmodulin-dependent protein kinase IV kinase isoform, Cam-kinase kinase alpha"

107..1624

/gene="Ca2+/calmodulin-dependent protein kinase IV kinase isoform, Cam-kinase kinase alpha"

/note="This sequence comes from Fig. 1; CaM-kinase kinase alpha"

/codon_start=1

/product="Ca2+/calmodulin-dependent protein kinase IV kinase isoform"

/protein_id="AAB46910.1"

/db_xref="GI:1836161"

/translation="MERSPAVCCDDPAELVERVAALVSAHLEAEPEGPEASGVDP PPARASVIFGASRFTVPRLSARSLQBRPAGSLCAQVGYSTGPASHMSR AMRPTIESHVAISDTECVQLNQYLQSEIGKAYGVRLAYNREDRHYAMKVL KKKLKOYGPFRPPPPGQAPQGGAKQLPLERVQETAILKKLDHVNVLIEVL DDPADNLXVLVDFLLRKGPVWVEPCDKPFPEQARLYLRDIIILGLEYLHCOQIVHRDI

ORIGIN

Query Match 81.7%; Score 1238.2; DB 10; Length 3429;

Best Local Similarity 88.6%; Pred. No. 3.3e-257;

Matches 1342; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

Qy 1 ATGAGAGGGGTCAGCTGTCTGTGCGAGGATCTCTGGGAGAGACTGGTAGAAACGGGTG 60

Db 107 ATGAGAGCGATCGAGCGCTCTGTGCGCAAGACCTCGAGCGGAACCTGGTAGAGCGGTG 166

Qy 61 GCAGGCATCGATGTGACTCACTTGGAGGAGCAGATGCTGGCCACAGAGCTACTAGAAAC 120

Db 167 GCAGGCATCGATGTGCGCCACCTAGAGAGGAGAGGAGGCGCCAGAGCTTCCAGCAAT 226

Qy 121 GGTGTGGAACCCCAACACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180

Db 227 GGTGTGGAACCCCAACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 286

Qy 181 CTGTCTCCAGCCCGGCTAGCTCTCAGCAGGAGAGCTTTCCTTACAGAGAGGGCCAGCA 240

Db 287 CCCACACCAAGCTGCTCCAGCGCTCTCTGTAGAAAGTTCTCCCTCGAGAAAGACGAGCG 346

Qy 241 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGTCAGGGGCTTCCAGCAGCACAATCTCC 300

Db 347 GGAAGCTGTCTAGAGGCTCAGGTGGGCTTACTCTACAGGACCTTCCAGTACATGTCT 406

Qy 301 CCCCAGGCTGTGGGAGGCGCCACCATCGAGTCCCACCGAGTGGCCATCTCAGATGAGAG 360

Db 407 CCTCGGCTTGGCGGAGACCCACCATCGAGTCCCACCATGTGGCCATCTCAGACACAGAG 466

Qy 361 GACTCGCTGACGCTGAACAGTACAGCTGCAGAGTGAGATTGGCAAGGCTGCTACGCT 420

Db 467 GACTCGTGCACACTGAACCAAGTACCAAGCTGCAGAGTGAGATTGGCAAGGCTGCTATG 526

Qy 421 GTGTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 480

Db 527 GTGTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 586

Qy 481 AAAAAAGAGTTTACTGAAGCAGTATGGCTTTCACGTGCGCCCTCCCGGAGAGGTCCTCCAG 540

Db 587 AAAAAAGAGTTTACTGAAGCAGTATGGCTTTCCTCGCCGCTCCTCCCGGAGAGGTCCTCCAA 646

Qy 541 GCTGCCAGGAGGAGCCACCAAGCAGCTGCTGCCCTCGAGCGGTGTACCAAGAGAGATT 600

Db 647 GCTCCTCAGGAGGAGGAGCCAAACAGCTGCTGCCCTCGAGCGGTGTATACCAAGAGAGATT 706

Qy 601 GCATCCTGAAGAAGCTGGACCAACGTAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660

Db 707 GCATTTTAAAGAACTGGACCAACGTAATGTAGTCAAGTTGATCGAGGTCTCTGGATGAT 766

Qy 661 CCAGCTGAGGAGCAACCTCTATTGGTGGTGGCTCTCTCAGAAAGGGGCGCTCATGGAA 720

Db 767 CCTGCTGAAGACAATCTCTATTGGTGGTGGCTCTCTCAGAAAGGGGACGATCATGGAA 826

Qy 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCGCTCTACCTGGGGACGTCATC 780

Db 827 GTGCCCTGTGACAAGCCCTTCCAGAGGAGCAAGCTGCGCTCTACCTCGGGACATCATC 886

Qy 781 CTGGGCTCTGAGTACTTGCATCTGCCAGAAAGATCGTCCACAGGAGATCAAGCATCAAC 840

Db 887 CTGGGCTCTGAGTACTTGCATCTGCCAGAAAGATGTGCACAGGAGATCAAGCGCTCAAT 946

Qy 841 CTGCTCTGGGAGATGATGGGACGTAAGATGCGGACTTTCGCTTCACCAACCACTTT 900

Db 947 CTGCTCTGGGAGCAGTGGGACGTAAGATGCGGACTTTCGCTTCACCAACCACTTT 1006

Qy 901 GAGGGGAACACACCTCAGCTGTCCAGCAGCGGGGAACCCAGCAATTTCATGGCCCCCGAG 960

Db 1007 GAGGGGAATGATCTCAGCTGTCCAGTACCGAGTACCGAGGAGCCCGCAGCATTCATGGCCCCCGAG 1066

```

961 GCCATTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCACTGGC 1020
Dbb|||||
1067 GCCATTCTGACACCGGCCAGAGCTTCAGTGGGAAGSCCTTGGATGTATGGGCCACTGGG 1126
QY|||||
1021 GTCAGTGTGACTGCTTGTCTATAGGAAGTGCCCATTCATCGACGATTCATCTCGGCC 1080
Dbb|||||
1127 GTTACACTGTATTGTTGTCTAGCGGAAGTGCCCATTCATTCATGATGATCATCTCGGCC 1186
QY|||||
1081 CTCACAGGAAGATCAAGATAGCCCGTGTGTTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Dbb|||||
1187 CTCACAGGAAGATCAAGATAGCCCGTGTGTTCTCTGAGGAGCCAGAAATCAGCGAG 1246
QY|||||
1141 GAGCTCAAGACCTGATCTCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1200
Dbb|||||
1247 GAACTCAAGACCTGATCTCTGAAGATGCTAGACAAGAAATCCTGAAACAAGATGGGGTG 1306
QY|||||
1201 CAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGAGAGAGCCCTTCCTCTGGAG 1260
Dbb|||||
1307 TCTGATATCAAGTTACACCCCTTGGTGACCAAGCATGGAGAGAGCCCTTCCTCTCAGAG 1366
QY|||||
1261 GAGGAGCACTGCAGCGTGTGGAGGTACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Dbb|||||
1367 GAGGAACATGCACTGTGTTAGAGGTGACTGAGGAGGAGTGAGAACTCAGTCAAGCTC 1426
QY|||||
1321 ATCCCCAGCTGGACCAACGCGTGATCTCTGGTGAAGTCCATGCTGAGGAAGCTTCTTTGGG 1380
Dbb|||||
1427 ATCCCCAGCTGGACCACTGTGATCTCTGGTCAAGTCTATGCTGAGAAAGCGTTCCTTTGA 1486
QY|||||
1381 AACCCGTTTGAGCCCCCAGGACCGGAGGAGAGCGATCCATGCTGTCAGAAACCTTA 1440
Dbb|||||
1487 AACCCATTTGAGCCCCCAGGACCAAGCAAGGAGAAAGATCTATGCTGCGCCAGGAACTTA 1546
QY|||||
1441 CTGGTGAAGAAGGGTTTGTGTAAGGGGCAAGAGCCAGAGTCCCGCGCTCCAGGAA 1500
Dbb|||||
1547 CTGTTGAAGAGAGATGTGGAGAGGGGCAAGAGCCAGAGCTTCGCGAGTCCAGGAA 1606
QY|||||
1501 GAGGAGGCTGCATCC 1515
Dbb|||||
1607 GATGAGGCTGCATCC 1621

RESULT 15
AB023658
LOCUS
DEFINITION
Rattus norvegicus mRNA for Ca/calmodulin-dependent protein kinase
kinase alpha, CaM-kinase kinase alpha, complete cds.
ACCESSION
AB023658
VERSION
AB023658.1 GI:4512333
KEYWORDS
Ca/calmodulin-dependent protein kinase kinase alpha, CaM-kinase
kinase alpha.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (sites)
Okuno,S., Kitani,T. and Fujisawa,H.
Evidence for the existence of Ca2+/calmodulin-dependent protein
kinase IV kinase isoforms in rat brain
J. Biochem. 119 (6), 1176-1181 (1996)
MEDLINE
96425004
PubMed
8827455
REFERENCE
2 (bases 1 to 3411)
Fujisawa,H. and Okuno,S.
Direct Submission
Submitted (15-FEB-1999) Hitoshi Fujisawa, Asahikawa Medical
College, Department of Biochemistry; Nishikagura 4-5-3-11,
Asahikawa 078-8510, Japan (E-mail:Okuno@asahikawa-med.ac.jp,
Tel:+81-166-68-2340, Fax:+81-166-68-2349)
FEATURES
Location/Qualifiers
1..3411
/organism="Rattus norvegicus"
source
```

```

/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain"
/tissue_lib="brain CDNA"
107..1624
CDS
/codon_start=1
/product="Ca/calmodulin-dependent protein kinase kinase
alpha, CaM-kinase kinase alpha"
/protein_id="BAA75246.1"
/db_xref="GI:4512334"
/translation="MERSPAVCCDDPRAELVERVAASVAHLEEBEGPPEPANGVDP
PPRAAASVPGSASPTVPRLSLKSFSLQRPAGSCLEAQVGYSTPAGSHMSPR
ANRPTIESHVAISDTECVQLQSEIKLGAVGVVRLAYNREDRHYAKVLIS
KKLLKQYGFPRRPPRPGSOAPOGPAKOLLPLERVYQETAIILKLDHNVNKLIVL
DIPAEQNLVLFVLLKQPMVEPCDKPPEQARLYLRDILLGLYLHCQKLVHRDI
KSNLLGLDDGHVKIADFVGSQFEGNDAQLSSTAGTFAPFAPEALISDTGQSFSGKAL
DWATQVTLFCFYVKGKPFIDEXYLAHHRKIKNEAVFPPEPEVSEBELKDLILKLDK
NPETRIGVSDIKLHPWVTKHGEPLFSEBEHCSVVEVTEEVKNVSKLIPSWTTLIV
KMLRKRISFGNPFEPQARREERSMSPGNLLLLKEGCGEGKSPELPGVQDEAAS"
3411
/note="18 a nucleotides"
polyA_site
ORIGIN
Query Match 81.6%; Score 1236.6; DB 10; Length 3411;
Best Local Similarity 88.5%; Pred. No. 7,4e-257;
Matches 1341; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCCTCGGGCAGAGCTGTAGAACGGTG 60
Dbb|||||
107 ATGGAGCGCAGTCCAGCCCTGTCTGCCAAGACCCCTCGAGCGGAACCTGTAGACGGGTG 166
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTTGGCCCGCAGAGCCTACTAGAAAC 120
Dbb|||||
167 GCAGCCATCAGTGTGCGCCACCTAGAGAGGAGAGAGGGCCAGAGCTGCCGCAAT 226
QY 121 GGTGTGGAACCCCGCCACCGGGCCAGAGTGCCTCTGTGATCCTCGCAGTACTTCAAGA 180
Dbb|||||
227 GGTGTGGAACCCCGCCACCGGGCCAGAGTGCCTCTGTGATCCTCGCAGTGCCTCAAGA 286
QY 181 CTGCTCCAGCGCGGCTAGCCTCTCAGCCAGGAGCTTCCCTACGAGCGGCCAGCA 240
Dbb|||||
287 CCACACCAAGTGGTCCCAGCCCTCTCTGTAGAAAGTCTCTCCCTGCAGGAAACACAGCG 346
QY 241 GGAAGTATCTCGAGCGCGCAGGCTGGGCGCTTATGCCACGGGGCTCCCGACCATCTCC 300
Dbb|||||
347 GGAAGTGTCTAGAGCTCAGGTGTGGGCCCTTACTCTACAGGACCTGCCAGTCACTGTCT 406
QY 301 CCCCAGGCTGGCGGAGGCCACCACTCGAGTCCCACCGTGGCCATCTCAGATGACAGAG 360
Dbb|||||
407 CCTCGGCTTGGCGGAGAGCCACCATCGAGTCCACCATGTGGCCATCTCAGACACAGAG 466
QY 361 GACTGCTGACGCTGNAACGAGTACAGCTGCAGATGAGATTGGCAAGGGTGCCTAGGTT 420
Dbb|||||
467 GACTGTGTCAACTGAACCAAGTACAGCTGCAGATGAGATTGGCAAGGGTGCCTATGTT 526
QY 421 GTGGTGAGGCTGGCCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTCC 480
Dbb|||||
527 GTGGTGAGGCTGGCCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTCC 586
QY 481 AAAAAGAAAGTACTGAAGCAGTATGCGCTTTTCAACGTCGCGCTCCCCCGAGAGGGTCCAG 540
Dbb|||||
587 AAAAAGAAAGTACTGAAGCAGTATGCGCTTTTCTCGCGTCTCCCCCGAGAGGGTCCCAA 646
QY 541 GCTGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGCGGGTGTACACGAGATT 600
Dbb|||||
647 GCTCCTCAGGAGGGCCAGCCAAACAGCTGCTGCCCTGGAGCGGTGTATACGAGGAGATT 706
QY 601 GCCATCCTCAAGAAAGCTGGACCACTGTAATGTGGTCAAACTCATCGAGTCTCTGGATGAC 660
Dbb|||||
707 GCATTTTAAAGAAACTGGACCACTGTAATGTAGTCAAGTGTCAAGTCTCTGGATGAT 766
QY 661 CCAGCTGAGGACCAACTCTATTGTTGTTGATCTCTCGAGAAAGGGGCGGCTCATGGAA 720
Dbb|||||
```

```

Db      767 CCTGCTGAAGACAATCTCTATTGTTGGTGTGACCTCTGAGAAAGGACCAGTCAATGGAA 826
QY      721 GTGCCCTCTGACAAGCCCTTCTGGAGAGAGCAAGCTCGCTCTACTGCGGGACGTCAATC 780
Db      827 GTGCCCTCTGACAAGCCCTTCCAGAGAGCAAGCTCGCTCTACTGCGGGACATCATC 886
QY      781 CTGGGCTCTGAGTACTTGCACCTGCCAGAGATCGTCCACAGGACATCAAGCCATCCAAC 840
Db      887 CTGGGCTCTGAGTACTTGCACCTGCCAGAGATTTGCAAGGACATCAAGCCGTCCAAT 946
QY      841 CTGCTCTCTGGGGATGATGGGCACTGAGATCGCGACTTTGGCGTCAAGCAACCAAGTTT 900
Db      947 CTGCTCTCTGGGGACGATGGGCACTGAGATCGCGACTTTGGTGTGAGCAACCAAGTTT 1006
QY      901 GAGGGGAAGCGCTCAGCTGTCCAGCAGCGGGGGAACCCGAGCAATTCATGSCCCCCGAG 960
Db      1007 GAGGGGAATGATGCTCAGCTGTCCAGTACGGCAGGACCCGAGCAATTCATGSCCCCCGAG 1066
QY      961 GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGC 1020
Db      1067 GCCATCTCTGACACCGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGG 1126
QY      1021 CTCACGTTGTAAGTCTGTTGATGGGAAGTCCCAATTCATCGACGATTTCACTCTGGCC 1080
Db      1127 GTTACACTGTAATGTTTGTCTACGGGAAGTSCCAATTCATTGATGAGTACATCTCTGGCC 1186
QY      1081 CTCCACAGGAAGATCAAGATGAGCCCTGGTGTTCCTGAGGAGCCAGAAATCAAGCGAG 1140
Db      1187 CTGCAAGGAAGATCAAGATGAGCCCTGGTGTTCCTGAGGAGCCGAGGTCAAGCGAG 1246
QY      1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTCGGGTG 1200
Db      1247 GAATCTAAGACCTGATCTGAAGATGCTAGACAAGATCCCTGAACAAGAAATTCGGGTG 1306
QY      1201 CCAGACATCAAGTTGCACCTTTGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1260
Db      1307 TCTGATATCAAGTTACACCTTTGGGTGACCAAGCATGGAGAGGAGCCCTTCCTTCAGAG 1366
QY      1261 GAGGAGCACTGAGCGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db      1367 GAGGAACACTGCACTGTGTGAGGTGACTGAGGAGGAGGTGAAGAACTCAGTCAAGCTC 1426
QY      1321 ATCCCCAGCTGGACCAAGGTGATCCTGTGAAGTCCATGCTGAGGAAGGTTCCCTTTGGG 1380
Db      1427 ATCCCCAGCTGACCACTGTGATCCTGTCAAGTCTATGCTGAGAAAGGTTCCCTTTGA 1486
QY      1381 AACCCGTTTGAGCCCGCAGGACGAGGGAAGAGCGATCCATGCTGCTCCAGGAACCTTA 1440
Db      1487 AACCCATTTGAGCCCGCAGGACGAGGGAAGAGATCTATGCTGCGCCAGGAACCTTA 1546
QY      1441 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGGCCAGAGCTCCCGGCGTCCAGGAA 1500
Db      1547 CTGTTGAAGAGAGGATGTGGGAAGAGGGGCAAGGCCAGAGCTTCCCGGAGTCCAGGAA 1606
QY      1501 GACGAGGCTGCATCC 1515
Db      1607 GATGAGGCTGCATCC 1621

```

Search completed: July 25, 2004, 09:54:21
 Job time : 9539 secs